

# 늪지인간과 지향성\*†

신상규‡

표상, 의미, 혹은 지향성의 개념을 자연화하기 위한 시도 중의 하나인 목적론적 이론의 난점을 지적하기 위한 논변으로 "늪지 인간(swampman)" 혹은 "우연적 복제물(accidental double)" 논변이 있다. 늪지인간은 우연적인 사건에 의해 생겨난 어떤 사람의 완전한 물리적 복제자이다. 목적론적 의미론에 따르면 우리는 이 복제인간(늪지인간)에게 어떠한 지향적 상태도 귀속시킬 수 없다. 많은 철학자들은 이러한 결론이 대단히 반직관적이라고 생각한다. 이 논문의 목적은 목적론자들의 이러한 주장을 가능한 한 옹호하기 위한 것이다. 논문은 크게 두 부분으로 나뉘어져 있다. 먼저 그러한 반직관성의 상당 부분이 지향적 상태에 대한 내재주의적 입장에 근거하고 있다고 보고, 외재주의와 내재주의의 구분에 대한 논의를 통하여 이러한 반직관성을 경감시키려 할 것이다. 두 번째로 '지향적 상태'가 일종의 '규범적인' 현상이라는 것을 전제했을 때에, 역사성을 요구하는 목적론적 의미론이 온건한 외재주의를 포함하는 여타의 자연주의적 이론에 비교했을 때에 갖게 되는 강점을 보이려 할 것이다. 이는 목적론의 주장이, 지향적 태도가 갖는 규범성을 확보하기 위한 가장 유망한 설명이라는 의미에서, 지향적 태도의 개별화에 대한 일종의 최상의 설명으로의 주리(an inference to the best explanation)라는 것을 지지하기 위한 것이다. 이런 관점에서 보면 늪지 인간에 대한 목적론의 대응은 그렇게 반직관적이기만 한 것은 아니다.

**주 제** 심리철학, 언어철학  
**주 요 어** 지향성, 지향성의 자연화, 지향성의 규범성, 목적론적 의미론, 외재주의, 늪지인간

\* 접수완료: 2003. 6. 2. / 심사 및 수정완료: 2003. 9. 24.  
 † 이 논문의 초고를 읽고 적합한 비판을 통하여 많은 부분이 수정 보완할 수 있도록 도와주신 익명의 심사위원들께 감사드립니다.  
 ‡ 서강대 강사

데이빗슨에서 예를 따온 ‘높지인간’의 논변은 다음과 같이 진행된다.<sup>1)</sup> 어느 날 갈대가 우거진 늪지의 나무 위에 벼락이 떨어졌다. 엄청난 전류가 흐르면서 그 옆에 서있던 나의 몸은 분자단위의 구성요소로 분해되고, 나무의 분자들이 절묘하게 결합함으로써 물리적, 기능적인 면에서 나와 완전히 똑 같은 높지인간이 생겨났다. 이 높지 인간은 나와 수적으로 (numerically) 구분될 뿐만 아니라, 그 구성 분자들도 사실은 수적으로 다른 요소들로 이루어져 있다. 하지만, 현재의(current) 인과적, 성향적 속성의 차원에서 이 높지인간은 나와 별반 차이가 없어 보인다. 사람들을 만나고, 이야기하고, 심지어 사랑도 하는 것처럼 보인다. 말하자면 이 높지인간은 시스템의 내부적인 물리적 구성과 그 상태, 그리고 관찰 가능한 외향적인 움직임에 있어서 나와 완전히 일치 한다.

전통적인 기능주의 혹은 행동주의적 분석에 따르면 이 높지 인간은 나와 완전히 동일한 심리적 상태를 가지게 될 것이다. 전통적인 기능주의에서는, 비록 심리적 상태가 두뇌의 상태와 동일시되지 않는다 하더라도, 기능적 상태로서의 심리적 상태는 여전히 우리의 신체나 두뇌의 상태에 수반하는 것이다. 그러므로 이 높지 인간이 나와 완전히 똑 같은 신체적, 두뇌적인 상태를 가지고 있다면, 이것이 가지고 있는 기능적 상태 또한 나와 완전히 일치할 것이다. 어떤 대상에게 심리적 혹은 지향적 상태를 귀속시키는 인식론적 기준으로서의 튜링 테스트를 생각하여도 동일한 결론이 도출된다. 높지인간은 튜링 테스트를 완벽히 통과할 것이다. 그렇다면 우리는 이 높지 인간이 지향적인 태도로서의 심리적인 상태를 가진다고 말할 수 있는가? 가령 이 높지 인간에게 어떤 믿음, 혹은 욕구를 귀속시킬 수 있는가? 당신의 직관은 어떠한가? 어떤 측면에서 높지인간 논변은 서얼의 중국어 방 논변에 비교될 수 있다. 서얼의 중국어 방

---

1) Davidson (1987), p.19.

논변이 보여주고자 했던 것은 구문(syntax)과 의미(meaning)를 구분하고, 구문만으로는 의미를 귀속시킬 수 없다는 것이었다. 서얼은 이러한 질문을 통하여 의미를 구성하는 것이 무엇인지에 대한 질문을 던진다. 지향적 태도란 기본적으로 의미 혹은 내용(content)을 갖는 상태이다. 높지인간 논변은 서얼적인 의미에서의 구문을 포함할 뿐만 아니라 더욱 포괄적인(내적인) 상태를 높지인간이 갖는다고 인정한 후에, 높지인간에게 의미 혹은 내용을 갖는 지향적 태도를 귀속시킬 수 있는가 하고 묻는 것이다.

목적론적 의미론(teleosemantics)에 따르면, 지향적 상태가 본질적으로 갖게 되는 내용(content)은 그 지향적 상태의 고유기능을 통하여 결정된다. 이들이 말하는 고유기능이란 진화 생물학 등에서 말하는 일종의 발생론적(etiological) 개념으로서, 지향적 상태 혹은 내용의 자연화와 관련하여 크게 두 가지의 역할을 담당하게 된다. 첫째는 지향적 내용이 갖는 의미의 규범성의 자연화이다.<sup>2)</sup> 크립키에 의하면 의미란 기본적으로 기술적(descriptive)인 것이 아니라 규제적 혹은 규범적인 범주에 속하는 것이다. 지향적 의미는, 내가 거기에 (우연히) 사실상 부합하는 행동적 성향을 갖는다는 것이 아니라, 마땅히 내가 그렇게 따라야만 하는(ought to, supposed to) 옳음 내지 올바름의 표준이라는 규칙으로서의 지위를 누린다. 고유기능의 개념은, 발생론적으로 결정되는 생물학적인 정상성과 같은 목적적 개념을 통하여, 지향적 내용이 누리고 있는 고유한 규칙 따르기로서의 규범적 성격을 자연화하기 위한 유망한 전략을 제시한다. 둘째로, 고유기능의 개념은 지향적 상태가 갖는 내용에 관한 여러 경쟁적인 해석들 중에서 어떠한 해석이 표준적인 해석인지를 집어내는 원리적인 절차에 관한 답변을 제공한다.<sup>3)</sup> 가령 밀리칸에 의하면, 믿음 표상의 내용은 그 표상의 소비자(혹은 소비하는 장치)가 정상적인 설명(normal explanation)에 부합하는 방식으로 그것의 고유한 기능(proper function)을

2) 의미론적 규범의 자연화와 관련한 목적론적 입장에 대한 자세한 논의는 신상규 (2003a)를 참조하라.

3) 고유기능을 통한 내용의 고정에 대한 자세한 논의는 신상규 (2003b)를 참조하라.

수행하기 위하여 전제하여야만 하는 정상적 조건(normal condition)에 의하여 결정된다. 이러한 절차를 통하여 앞서 언급된 올바른의 표준으로서의 지향적 내용이 결정되며, 이러한 표준에서 일탈한 심적 표상의 발생이 바로 오표상의 사례가 된다.

그런데, 목적론적 의미론자들이 의도하는 고유기능의 개념은 그 구성적 본질에 있어서 발생론적, 즉 역사적인 것이다. 어떤 대상이 고유의 기능을 갖는다는 것은 가령 진화의 역사적인 과정 속에서 그 대상이 그렇게 기능함에 대한 실제적인(actual) 선택이 이루어졌음을 의미한다. 거칠게 정의하자면, 어떤 유기체 O가 가지고 있는 기관 X가 Y라는 고유한 기능을 갖는다는 것은, 과거 O의 조상들에서 X와 같은 종류의 것들이 Y를 수행함으로써 O가 속하는 종족의 생존 및 보존에 기여하였고, 그 결과 X가 현재의 형태로 O에 남아있다는 것이다. 인간이나 동물이 갖는 표상 시스템도 그것이 고유 기능적 범주에 속하는 한에 있어서 예외일 수 없으며, 고유 기능을 획득하기 위해서는 그에 걸맞는 적절한 선택의 역사를 가지고 있어야 한다. 진화의 역사에 있어서 생물체가 어떤 새로운 특징을 가지게 되는 과정은 다양하다. 그 중 대표적인 것이 유전자 변이에 의한 것이다. 그러나 밀리칸과 같은 목적론자들에게는 이렇게 생겨난 1세대의 특징들은 어떠한 고유 기능도 갖지 못한다. 이런 특징들이 후속 세대를 거치면서 해당 생물체(종)을 위하여 모종의 역할을 수행하고 그 유용성을 인정받아서 안정화될 때만이 고유의 기능을 획득 하게 된다. 물론, 기능에 대한 발생론적 사실이 꼭 생물학적 진화에 한정될 필요는 없다. 가령 밀리칸이나 드레츠키 같은 이에게 있어서, 학습과정(learning process)이나 어떤 언어 공동체를 통하여 특징의 어휘가 계승되고 살아남는 것도 일종의 역사적 선택의 과정을 거치는 것이다. 그리고 인공물(artifact)의 경우에 있어서 적절한 선택의 역사는 디자이너의 의도에 의해서 주어진다.

그런데 만일 심적 표상과 같은 우리의 지향적 상태가 이들의 주장과 같이 그 본질에 있어서 고유기능적 범주에 속하는 것이라면, 그래서 그

것이 갖는 지향성 자체가 부분적으로는 과거의 역사에 의해 결정되는 것이라면, 우리의 지향적 상태는 단순히 우리의 두뇌가 가지고 있는 현재의 인과적 혹은 성향적 속성들만을 통하여 결정되지 않는다. 다시 말해서 지향적 상태는 우리 두뇌의 현재적인 물리적 상태에 수반하지 않는다. 그런데, 앞에서 예로 든 높지 인간은 어느 순간 발생했던 우주적 우연의 산물일 뿐이며, 보통의 우리네 인간이 가지고 있는 탄생등과 관련한 적절한 선택의 역사를 가지고 있지 않다. 목적론자들의 주장처럼, 만일 믿음이나 욕구와 같은 표상 상태가 고유기능(proper function)의 범주에 속하고 고유기능의 개념이 그 본질에 있어서 역사적 개념이라면, 비록 높지 인간이 현재의 물리적 상태나 관찰 가능한 행동적 성향으로 보아서 나와 전혀 구분이 되지 않는다 하더라도 우리는 믿음이나 욕구와 같은 지향적 상태를 그에게 귀속시킬 수 없다. 많은 철학자들은 높지인간에게 아무런 지향적 상태도 귀속시킬 수 없다는 이러한 결론이 대단히 반 직관적인 것이라고 생각한다.

고유기능을 통한 지향적 태도의 분석이라는 목적론적 의미론의 기본적인 주장을 유지하면서도 높지 인간에게 모종의 지향적 태도를 귀속시키려는 시도가 있을 수 있다. 그 첫 번째로 우선 높지인간에게 모종의 역사성을 부여하려는 시도가 있을 수 있다. 가령, 앞에서 예로 든 높지 인간은 나의 복제물이다. 비록 생식과정이라는 정상적인 과정을 거치지지는 않았지만 여전히 나의 DNA가 그에게 복사되었으며, 그런 한에 있어서 나의 DNA에 내재되어 있는 유전적인 역사를 그 높지인간도 공유하고 있다고 주장할 수 있을 것이다. 정상적인 탄생의 경우에도 인간이 가지고 있는 생물학적 진화의 역사는 결국 유전자를 통한 복제를 통해 계승되는 것이므로, 높지 인간이 보통의 인간과 다른 점은 단지 어떠한 메커니즘을 통하여 유전자 복사가 이루어졌느냐의 문제일 뿐이라는 것이다. 그러나 이러한 고려는 나로부터 복제되지 않은, 정말로 우주적 우연에 의해 생겨난 높지인간을 상상할 수 있으므로, 그 근본적인 해결책이 될 수 없다. 다시 말해서, 나와 아무런 연관이 없이 지구의 반대편에 나와

똑같은 늑지 인간이 우연적으로 탄생하였다고 가정해보자. 이 존재는 나와 동일한 구조를 가지고 있고, 나의 유전자에 대응하는 모종의 '세포'구조도 가지고 있을 것이다. 그런데, 이 존재는 말 그대로 나오는 아무런 상관없이 우연적으로 생겨난 존재이므로, 아무리 그의 '유전자'가 나의 유전자와 동일하다고 하더라도, 여기에 어떤 역사성을 부여할 수는 없다.

두 번째 가능한 목적론적 시도는, 생물학적 기능이란 것을 비역사적으로 이해하는 방식이 될 것이다. 이러한 시도는 가령 우리가 뒤에서 간단히 살펴보게 될 커밍스의 기능적 분석 같은 것을 활용하는 방식이다. 그런데, 앞에서 지적하였듯이, 목적론자들이 기능개념을 끌어들이는 이유는 이것이 기술적 성향에 대비되는 것으로서의 규제적인 규범 개념을 자연화할 수 있는 유망한 전략을 제시하기 때문이다. 하지만 밀리칸이나 니앤더(Neander) 같은 목적론자들은 지향적 규범성은 고유기능에 대한 역사적 이해를 통해서만 확보될 수 있다고 믿는다. 즉, 커밍스적 기능 분석은 일종의 성향적 기능분석으로서, 이것으로부터 고유한 목적을 갖는 것으로서의 규범적 기능개념을 끌어낼 수가 없다는 것이다. 그래서 이들은 일종의 양립논제를 추구하기 보다는 늑지 인간이 아무런 지향적 상태를 갖는 않는다는 반직관적인 결론을 기꺼이 받아들인다.

이 논문의 기본적인 목표는 늑지 인간에게 지향성을 귀속시킬 수 없다는 목적론자들의 주장이 처음에 들리는 것처럼 그렇게 반직관적이기만 한 것은 아니라는 점을 보이는 데에 있다. 이러한 논의는 두 단계로 나누어 진행될 것이다. 먼저 늑지인간이 지향적 상태를 갖지 않는다는 것이 반직관적으로 들리는 이유 중의 많은 부분은, 지향적 상태에 대하여 우리가 암암리에 일종의 내재주의(internalism)를 전제하기 때문이라는 점을 논의할 것이다. 지향적 상태란 본질적으로 그것이 갖는 내용에 의하여 개별화되는 상태들이다. 개별화의 문제와 관련지어서 생각해 본다면, 내재주의와 외재주의(externalism)의 구분은 늑지인간과 관련된 반직관성을 이해하는데 매우 유용한 분석의 틀을 제공한다. 목적론적 의미론은 기본적으로 외재주의적인 입장이다. 우리가 첫 번째 할 일은 외재주의를

옹호함으로써, 높지인간에 대한 목적론적 주장이 갖는 반직관성을 가능한 한 완화시켜 보는 것이다. 그런데, 외재주의가 옹호된다고 해서 바로 목적론적 주장이 정당화 되는 것은 물론 아니다. 내재주의자에서 외재주의자로 이어지는 일종의 스펙트럼을 생각해 본다면, 높지인간에 대한 직관이 서로 나뉘는 부분은 외재주의자들이 서 있는 스펙트럼의 그 어디 중간 정도가 될 것이다. 말하자면, 내재주의자 전부와 외재주의자의 일부는 높지인간에게 지향적 태도를 귀속시키려 할 것이며, 목적론자와 같은 보다 강성의 외재주의자는 지향적 태도의 귀속을 거부하는 식이 될 것이다. 하지만, 높지인간이 아무런 지향적 태도를 갖지 않는다는 주장의 반직관성은 많은 부분이 내재주의적 틀에 근거한다.

논문의 두 번째 단계에서는, 크립키가 지적한 데로 지향적 내용이 기본적으로 규범적인 현상이라는 것을 가정하고, 이러한 규범적 성격을 설명하기 위해서 어떠한 조건들이 지향적 내용의 본질적인 구성적 조건으로 포함되어야 하는가를 물을 것이다. 이 과정에서 필자의 기본적인 논점은 매우 강력한 형태의 외재주의적 입장인 목적론적 의미론이 규범성에 대한 가장 설득력 있는 설명을 제공한다는 것이다. 즉, 지향성의 규범성을 적절히 해명하기 위해서는, 내적인 인과적 성향이나 공간적인 외부 환경 뿐만 아니라 시간적인 차원의 역사적 사실까지도 포함되어야 함을 주장할 것이다. 이때 비교의 대상이 되는 입장은 내재주의적인 입장뿐만 아니라, 역사성을 요구하지 않는 온건한 외재주의의 입장이 그 주 타깃이 될 것이다.

그런데 여기서 지향성이 기본적으로 규범적인 현상이라는 것을 전제한다는 말은, 우리가 적절한 규범성을 확보하지 못하는 한에 있어서, 지향성의 실재성도 인정될 수 없다는 요구로 이해될 것이다. 가령, 지향적 의미라는 것이 근본적으로 불확정적인 것이라면, 크립키 방식의 규범성의 요구는 당연히 만족될 수 없을 것이다. 그런 한에 있어서, 그리고 규범성이 지향성의 본질적 구성요건이라면, 이는 단순한 의미의 불확정성을 뜻하는 것이 아니라, 의미의 실재성에 대한 부정으로 이어질 수밖에 없

다는 것이다. 그런데, 크립키가 말하는 규범성이라는 것이 과연 진실로 지향적 태도를 본질적으로 구성하는 요건인가에 대해서는 별도의 철학적 논란의 여지가 있다. 가령, 드레츠키는 최근의 논문에서 비록 지향적 태도가 규범적 적용을 받을 수는 있지만, 그렇다고 해서 그러한 규범적 적용을 받는다는 것이 과연 지향성을 구성(constitute)하는 본질적 조건이어야 하는가에 대해서는 회의적 태도를 취한다.<sup>4)</sup> 그러나 본 논문에서는 이 물음을 다루지 않을 것이다. 본 논문의 기본적인 문제의식은, 앞서 언급한대로 규범적 성격이 지향성을 구성하는 본질적 조건이라는 것을 가정한다면, 어떻게 그러한 규범성을 자연주의적으로 확보할 것인가의 문제이다. 그런 의미에서, 역사적 사실에 입각한 외재주의의 한 형태인 목적론이 의미 혹은 지향적 내용이 갖는 규범성을 가장 잘 확보해 줄 수 있는 이론이라는 본 논문의 주장은, 일종의 선택적 논변의 한 형태로서 최상의 설명으로의 추리(an inference to the best explanation)라는 구조를 취하고 있다.

이 논문에서 주장하는 범위에 대하여 몇 가지 제한을 두고자 한다. 먼저, 우리가 여기서 관심을 가지고 있는 이론들은 무엇보다도 자연주의적 이론들이다. 현재의 인과적 혹은 성향적인 속성, 환경과의 관계, 그리고 역사적 조건 같은 것들은 우리가 가지고 있는 경험적 과학에 비추어 큰 문제가 없는 일종의 자연주의적 조건들이다. 목적론의 상대로 고려될 경쟁이론들은 이러한 조건들에 기반한 자연주의적 이론들에 국한될 것이다. 그리고 이 논문은 목적론에 대한 전반적이며 궁극적인 옹호를 그 목표로 하고 있지 않다. 다시 말해서, 이 논문의 목표는 목적론적 의미론이 모든 면에 있어서 성공적인 프로그램이라는 식으로 옹호하려는 것이 아니며, 단지 지향성이 갖는 규범성의 해명이라는 측면에서 볼 때 목적론이 다른 자연주의적 이론들에 비해 유망한 이론이라는 것이다. 그리고 그런 맥락에서 이해할 때, 늑지인간이 아무런 지향적 태도를 갖지 않는

---

4) Dretske (2000), p.252.

다는 목적론의 주장은 처음 생각했던 것보다는 훨씬 온건한 주장으로 이해될 수 있다는 것이다.

2

먼저, 높지인간과 같은 사유실험이 철학적 이론을 평가하기 위한 논거로서 얼마만큼 적절한가에 대한 반성으로부터 논의를 시작하여 보자. 밀리칸(Millikan)같은 이는 지향성이나 고유기능에 대한 자신의 이론이, 믿음이나 욕구와 관련한 우리들의 일상적인 언어 사용에 대한 필요충분조건을 탐색하는 개념적 분석이 아니라, 일종의 이론적 정의(theoretical definition)임을 분명히 한다. 그러므로 자신의 이론에 대한 궁극적인 옹호는 우리들의 일상적인 언어 사용에 얼마나 잘 부합하느냐에 의해서가 아니라, 이러한 개념들을 통하여 구성되는 이론이 얼마만큼 다양한 현상들을 잘 설명해주느냐에 따른 이론의 설명력에 의해서라고 주장한다.

우리의 일상적인 직관과 어떤 이론의 귀결이 불일치하는 것은 철학이나 과학에서 매우 흔한 사례이다. 멀리 갈 것도 없이 지구가 태양의 주위를 돌고 있다는 것은 얼마나 반직관적인 주장인가? 많은 경우에 있어서 정교하고 설명력이 있는 이론과 우리의 일상적인 직관이 충돌할 때에, 포기해야 할 쪽은 결국 우리의 일상적인 직관이었다. 어떤 좋은 이론의 뒷받침이 없는 직관은 말 그대로의 직관일 뿐이다. 우리의 직관은 흔히 어떤 것의 겉으로 드러난 징표에 의존한다. 그러나 어떤 것이 모종의 성질 자체를 갖는다는 것과, 그 성질을 갖는가를 판단하기 위해 사용하는 여러 가지 징표를 보여주고 있다는 것은 엄연히 다른 문제이다. 금처럼 보이는 모든 것이 금은 아닌 것이다. 목적론자들이 생각할 때, 높지인간의 예가 보여주고 있는 것은 단지 높지 인간이 지향적 상태에 관련된 여러 가지 징표를 가지고 있다는 것뿐이다.

여기서 우리는 높지인간이 제기하고 있는 도전을, 지향적 상태를 지향적 상태로 만드는 것이 도대체 무엇인가를 묻는, 일종의 구성적(constitutive)

원리에 대한 질문으로 이해할 수 있다. 가령, 목적론자는 튜링 테스트가 지향성에 대한 구성적 원리라는 것을 부정할 것이다. 많은 경우에 있어서 비록 우리가 겉으로 드러난 행동을 통한 일종의 튜링테스트를 통하여 지향적 상태를 귀속시키기는 하지만, 그러한 행동적 특성이 곧 지향적 태도 자체를 구성하는 것은 아니라는 것이다. 그렇다면, 늑지인간의 논변이 목적론적 이론의 부당성에 대한 적절한 비판으로 인정받기 위해서는, 지향적 상태를 구성하는 것이 무엇인가에 대한 좋은 이론적 뒷받침이 있어야 할 것처럼 보인다. 그러한 이론의 뒷받침이 없는 직관은 말 그대로의 직관일 뿐이며, 비판적 논점으로서의 설득력은 그만큼 떨어질 것이기 때문이다. 늑지 인간의 지향적 상태와 관련된 두 상충하는 직관에 대한 적절한 평가는, 이 상반되는 직관을 서로 뒷받침하는 이론적 토대들에 대한 비교를 통해서만 적절히 수행될 수 있을 것이다.

그렇다면, 늑지인간이 아무런 지향적 상태를 갖지 않는다는 목적론적 의미론의 이론적 성격을 어떻게 정리할 수 있을까? 그리고 이와 상반되는 직관의 이론적 배경은 또 어떻게 정리할 수 있을까? 나는 목적론적 의미론이 지향적 상태에 대한 매우 강한 의미의 외재주의를 채택하고 있으며, 늑지인간과 관련하여 목적론에 상반되는 직관의 많은 부분은 지향적 상태에 대한 일종의 내재주의에 근거하고 있다고 생각한다. 앞서 간략히 언급하였듯이, 지향적 내용은 기본적으로 규범적 현상이라고 할 수 있다. 만일 그렇다면, 우리는 일단 늑지인간과 관련된 문제를, 외재주의와 내재주의 중 어떤 이론이 규범적 현상으로서의 지향적 상태를 설명하는 데 보다 적절한 이론이냐의 문제로 잠정적으로 재해석하고, 지향성의 구성원리로서의 내재주의가 가지고 있는 한계점을 지적함으로써, 늑지인간이 지향적 태도를 갖지 않는다는 목적론적 의미론의 주장을 부분적으로 옹호할 수 있을 것이다.

내재주의와 외재주의가 각기 어떤 입장이며 이들의 대립 관계를 어떻게 규정하는가에 대해서는 여러 가지의 다양한 관점이 있을 수 있겠지만, 이 논문에서는 맥긴(McGinn)의 정의를 따를 것이다.<sup>5)</sup> 먼저 내재주의

는 지향적 상태란 것이 환경과 분리되어 있는 주체에 관한 내부적 사실들을 통하여 결정된다는 입장이다. 다시 말해서 내재주의는 우리의 몸이나 두뇌 속의 내부적인 상태 혹은 사실들만으로 어떤 심적 상태를 규정하는데 충분하다는 입장이다. 물론 심적 상태와 관련한 이런 내적 사실들은 겉으로 관찰가능한 행동적 성향(behavioral disposition)이나 우리의 감각에 주어지는 직접적인 자극(proximate stimulation)을 포함한다. 여기서 중요한 점은 우리의 심적 상태를 규정하기 위하여 주체 바깥에 있는 외부적인 사실을 언급할 필요가 없다는 것이다. 흔히 이런 내부적 사실의 경계로 우리의 피부(skin)를 언급하기도 한다. 이런 내재주의에 기본적으로 깔려있는 생각은 심적 상태의 동일성에 대한 실체적(substantial) 관점이다. 말하자면 우리가 믿음이나 욕구와 관련한 심리적, 지향적 상태를 주체에 귀속시킬 때에, 우리는 이것들을 모종의 속성(property)을 실체로서의 주체에 귀속시키는 것으로 이해한다. 어떤 의미에서 이런 내재적인 접근 방식은 어떤 사물의 동일성을 판단하는데 있어서 우리가 디폴트(default) 값으로 채택하고 있는 조건이다. 우리가 기본적으로 사용하고 있는 언어는 그 성격에 있어서 일종의 사물중심 언어이다.

이에 반하여 외재주의는, 내가 가지고 있는 심리적 상태 혹은 지향적 상태를 규정하는데 있어서, 주체가 놓여있는 환경과 같은 외부적인 사실이 본질적인 구성원리로 작용한다는 입장이다. 다시 말해서 심리적 상태는 우리의 두뇌나 신체의 상태에 수반하지 않으며, 심리적 상태에 대한 수반의 기초 성질로서 우리로부터 독립적인 외부의 환경과 관련한 사실을 추가하여야 한다는 논제로 외재주의를 이해할 수 있을 것이다. 이러한 외재주의의 입장에서 지향적 상태는 외부적 환경과 관련된 일종의 관계적인 사실로 파악된다. 다시 말해서 마음은 세계에 대하여 근본적인 의미에서 자율성을 유지하고 있는 독립적인 영역이 아니다. 이런 의미에서 외재주의는 무엇이 나와 관련된 심리적 사실을 결정짓는가와 관련된

---

5) McGinn (1989), pp.1-9.

일종의 형이상학적 논제이며, 그 기본적인 주장은 외부의 세계가 나의 심리적인 상태를 결정하는 혹은 구성하는 일부분이라는 것이다. 그리고 기본적으로 이러한 결정성의 방향은 외부에서 내부로 향하여 있다. 그런 의미에서 외재주의는 상당히 강건한 외부세계에 대한 실재론을 전제로 한다.<sup>6)</sup>

높지 인간이 아무런 지향적 상태를 갖지 않는다는 것이 만일 반직관적으로 들린다면, 그 이유의 많은 부분은 지향적 상태라는 것이 높지인간이 가지고 있는 물리적인 상태나 행동적인 성향에 의하여 결정되는 것이라고 우리가 암암리에 전제하고 있기 때문이다. 높지인간의 예에서 우리가 인정하고 있는 것은 그것이 물리적인 측면에서 나와 완전히 동일하고

- 
- 6) 현재 논의되는 맥락에서의 내재주의는 퍼트남이 내적실재론(internal realism)에서 말하는 내재주의(internalism)와는 서로 다른 의미이다. 물론 밀리칸과 같은 목적론적 의미론자나 외재주의자들은 내적실재론의 대척점에 서있는 실재론자들이다. 밀리칸과 같은 강한 외재주의자들의 기본적인 주장은 세계가 마음을 규정한다는 것이다. 그런데, 우리가 외재주의적 주장의 대표적인 사례로 인용하게 될 쌍둥이 지구 논변을 내 놓은 사람은 퍼트남이다. 퍼트남의 경우, 비록 실재론에서 내적실재론으로 돌아섰다고 할지라도, 그가 “의미의 의미”에서 쌍둥이 지구 논변을 통하여 보여주었던 외재주의적 태도를 포기한 것 같지는 않다. (가령 「이성, 진리, 역사」의 “통속의 두뇌” 논변을 보라.) 그런데, 만일 외재주의가 실재론을 전제해야만 하는 것이라면, 이것은 밀리칸과 같은 사람에게 문제가 되는 것이 아니라 퍼트남의 문제가 된다. 퍼트남의 내적 실재론은, 마음과 세계가 서로를 규정한다는 주장을 통하여, 세계와 마음사이에 존재하는 일종의 지칭적 갭을 없애려는 시도로 볼 수 있다. 퍼트남은 마음과 세계가 서로를 규정한다고 말하고 있지만 그 강조점은 아무래도 마음에서 세계로 향하는 화살표에 놓여있는 것 같다. 맥킨은 상기한 책에서 이 문제를 “matching problem”이라고 부르는데, 그 문제의 근원은 개념(혹은 마음)과 대상(혹은 세계)을 서로 자율적이고 독립적인 것으로 보는데 있다고 진단한다. 지금 우리가 논하려고 하는 외재주의도 어떤 의미에서는 지칭(혹은 진리)의 가능성과 관련한 matching problem에 대한 한 해결책이다. 이들은 (퍼트남의 내적 실재론과는 반대의 방향에서) 마음이 세계를 통하여 규정된다는 주장을 통하여 이 문제를 해결하려 한다. 물론 외재주의자들도 마음과 관련한 모든 것이 외부적 사실만으로 규정된다고 주장하지는 않는다. 이들이 세계가 마음을 결정한다고 말할 때 의도하는 것은 부분적인(partial) 결정관계이다. 그런데 퍼트남이 쌍둥이 지구 논변의 외재주의적인 관점을 여전히 유지하기 위해서는 일종의 ‘내재적 외재주의(internal externalism)’를 주장해야만 하는 것처럼 보인다. 이런 입장이 어느 정도의 정합성을 가질 수 있을지는 별도의 논의가 있어야 할 것이다.

행동적인 측면에서 지향적 상태를 갖는 존재와 전혀 분간이 안 된다는 사실이다. 다시 말하자면, 여기서 우리가 인정하고 있는 것은, 위의 규정에 따를 때 높지인간과 관련된 내부적 사실이 나와 관련된 내부적 사실과 완전히 동일하다는 점이다. 그러므로 이런 전제로부터 지향적 상태라는 측면에서 보아 나와 높지인간이 크게 다른지 않다는 결론을 도출하는 가장 손쉬운 길은 모종의 내재주의적 입장을 전제하는 것이다. 높지인간이 지향적 태도를 갖지 않는다는 목적론적 주장에 대한 거부감의 많은 부분은 사실 지향적 태도의 구성원리로서의 외재주의에 대한 반발이라고 볼 수도 있다.

사실 이런 거부감은 지향적 심리학에 있어서 소위 좁은 내용(narrow content)이라는 개념을 통한 개별주의적 심리학을 주장하는 사람들의 시도와 맥이 닿아 있다. 그렇다면 이런 거부감을 뒷받침하고 있는 내재주의적 직관의 기본적인 논점은 무엇인가? 그 첫 번째 이유로 우리는 앞서 언급한 바 있는 실체적 사고를 들 수 있을 것이다. 물론 지향적 상태가 곧 실체적인 것은 아니다. 하지만 여기서 지향적 상태는 어디까지나 외부로부터 단절된 하나의 폐쇄된 체계로서의 실체적 대상이 갖는 속성적인 상태로서 이해되고 있다. 두 번째로 지향적 상태가 가지고 있는 인과적 효용성(causal efficacy)의 문제를 들 수 있다. 가령 전통적인 기능주의의 입장에 따르면, 하나의 심리적 상태는 그 심리 상태가 갖는 인과적 역할을 통하여 정의된다. 그런데 이 기능주의에서 말하는 인과적 역할이라는 것은 결국 따지고 보면 우리의 피부 안으로 제한되어 있는 인과적 역할이다. 그리고 그것은 앞서 언급한 바 있는 모종의 실체가 지니는 속성적인 의미에서의 인과적 역할이다.

김재권이 지적하듯이 어떤 것이 존재한다는 주장은 그것이 인과적 영향력을 갖는다는 것이다. 아무런 인과적 영향력이 없는 무엇이 존재한다는 주장은 별로 흥미로운 주장이 되지 못한다. 우리가 어떤 것을 그 어떤 것으로 만들어 주는 구성적인 원리가 여차여차한 것이라고 주장하고 있다고 가정해 보자. 이 때에, 그 구성적인 원리가 그 어떤 것이 가지고

있는 인과적 영향력을 해명해주어야 하는 원리여야 한다고 요구하는 것은 대단히 적절한 요구인 것처럼 보인다. 그런데, 우리가 인과적 영향력을 운운할 경우 그때의 인과성이란 개념은 대개의 경우 국소적인(local) 영향력의 문제이다. 다시 말해서 인과적 영향력의 주체로서 생각되는 대상이나 사건이 시공간적인 측면에서 근접해 있어야 한다는 의미에서 국소적으로 제한되며, 그 인과적인 영향력이란 것도 대상이나 사건이 가지는 내재적(intrinsic)인 속성을 통한 국소적인 영향력인 것이다. 만일 인과적인 영향력이 이러한 것이라면, 외재주의적으로 규정된 지향적 상태의 경우 어떻게 그 인과적 영향력을 행사할 수 있을 지가 대단히 모호해진다. 앞에서 우리는 이미 외재주의적으로 규정된 지향적 상태가 우리의 두뇌나 신체의 물리적 상태에 수반하지 않음을 지적하였다. 그런데 행위의 주체로서 우리가 행사하고 있는 인과적 영향력이 우리의 두뇌나 신체에 수반할 수밖에 없는 것이라면, 우리가 가지고 있는 하나의 상태로서의 지향적 상태가 가지고 있는 인과적 영향력이란 것도 결국에는 우리의 신체나 두뇌의 일부가 갖는 인과적 영향력에 수반할 수밖에 없을 것 같다. 그런데 지향적 상태라는 것이 외재적으로 규정된다면, 외부의 대상이나 역사적인 과거가 나의 행동을 일으키는 데에 모종의 별도의 인과적 영향력을 행사한다는 말인가? 만일 그렇다면 이는 인과적 영향력에 대한 대단히 이상한 이론으로 귀착된다.<sup>7)</sup>

7) 외재주의를 받아들일 때, 지향적 성질들이 갖는 인과적 효용성은 별도의 논의를 필요로 하는 것처럼 보인다. 참고로 이 문제에 대한 버지와 드레츠키의 응답을 소개한다. 버지는 지향적 태도에 대한 개별화의 원리로서의 외재주의와 인과성에 대한 국지주의는 서로 양립가능한 것이라고 논한다. (Burge (1986), p.16.) 외재주의의 논점은 두 가지 다른 환경이 행위 주체의 신체에 가하는 인과적 영향에 있어서는 서로 동일하지만, 행위 주체에게 전달되는 지향적 내용에 있어서는 차이를 만들 수 있다는 것이다. 즉, 지향적 상태를 외재적으로 규정한다고 해서, 행위의 발생과 관련하여 외부의 환경이 주체의 신체를 통하지 않는 어떤 별도의 인과적 영향력을 행사하고 있다고 생각할 필요는 없다. 환경이 문제가 되는 것은 지향적 태도의 개별화와 관련된 것이지, 지향적 태도가 갖는 인과적 힘에 영향력을 끼친다는 것은 아니라는 것이다. 그러나 이러한 주장에 대해서는 지향적 태도의 구획과 관련된 외부적 요소를 인과적으로 무력하게 만듦으로써 일종의 부산현상론에 빠진다는 비

여기서 우리는 내재주의가 국소적 인과성, 개별주의, 실체적 사고, 계산기계적 기능주의등과 같은 일련의 생각들과 서로 맞물린 채 하나의 견고한 틀을 형성하고 있음을 관찰할 수 있다. 그리고 높지인간이 지향적 상태를 갖지 않는다는 주장에 대한 거부감의 많은 부분은 바로 이러한 틀에 근거하고 있다. 자 그러면 높지인간에 대한 이런 주장이 결코 반직관적이지 않다는 것을 보여주기 위해서는 무엇을 하여야 하나? 여기서 가장 먼저 지적할 수 있는 것은, 외재주의자가 가지고 있는 기본적인 문제의식은 지향적 심리학에서 가정된 설명적 단위로서의 지향적 상태를 어떻게 구획(demarcation)짓는가의 문제에 대한 것이라는 점이다. 지향적 상태에 대한 제거주의적인 입장을 취하지 않는 한, 우리는 지향적 심리학을 모종의 실재하는 지향적 상태에 대한 경험과학으로 인정한다. 그리고 현재 받아들여지고 있는 표준적인 입장은 셀라스의 주장대로 지향적 태도를 지향적 심리학에서 가정하는 일종의 이론적 대상으로 간주하는 것이다. 이 경우, 우리가 어떤 대상에게 지향적 태도를 귀속시킨다고 할 때에, 하나의 이론적 대상으로서의 그러한 지향적 태도를 어떻게 적절히 개별화(individuate)할 것이냐의 문제가 자연스럽게 대두된다. 외재주의는

---

판이 적용될 수도 있다. 드레츠키는 국지적으로 문제가 되는 인과성과 지향적 설명의 맥락에서 문제가 되는 인과성을 구분함으로써 이러한 비판에 답하고자 한다. (Dretske (1987), p.160.) 가령 종을 울렸을 때(C) 조건화 되어있는 개는 침을 흘린다(E). 우리가 침을 흘리게 만든 원인이 무엇이라고 물었을 때, 종을 울린 것이나 이로부터 야기되는 모종의 청각적 경험 등은 국지적으로 문제가 되는 인과적인 힘으로서의 촉발적 원인에 해당하는 것이다. 드레츠키는 이러한 촉발적 원인을 통한 행위의 설명은 가령 신경생리학 등이 담당해야 할 문제라고 본다. 그런데, 개가 침을 흘리는 것이 단순히 신체적인 움직임이 아니고 행동인 한에 있어서, 우리가 지향적 설명의 문맥에서 해명하고자 하는 것은 이러한 촉발적 원인이 아니다. 지향적 설명에서 문제가 되는 것은, 종을 울렸을 때 개가 도망을 간다던가 하지 않고 왜 침을 흘리는가에 대한 이유를 묻는 것이다. 말하자면 지향적 설명에서의 인과성은 C를 가지고 E를 설명하려는 것이 아니라, 종을 울리는 사건으로부터 왜 침을 흘리는 행동의 패턴이 생겨났는가에 대한 질문, 즉 C->E 라는 구조가 생겨난 것에 대한 원인과 관련되어 있다는 것이다. 드레츠키는 이러한 원인을 구조적 원인(structuring cause)이라고 부른다. 그리고 이러한 구조적 원인을 해명함과 관련하여 외부적 환경이나 과거의 역사 등이 인과적 관련성을 띄게 된다. 이러한 주장에 대한 보다 포괄적인 논의는 Dretske (1988)을 참조하라.

지향적 태도를 개별화함에 있어서, 외부적 환경이 그 구성적 원리로 개입된다는 주장이다.

그런데 사실 이런 외재주의는 그 주장의 파격성에도 불구하고, 어떤 의미에서 우리에게 대단히 친숙한 논제이기도 하다. 그 대표적인 논의가 퍼트남의 쌍둥이 지구 논변이다. 쌍둥이 지구 논변의 핵심은 의미라는 것이 우리의 머리 속에 있는 것이 아니며, 어떤 단어의 외연은 기본적으로 화자가 처한 환경 속의 대상이 갖는 본질을 통하여 결정된다는 주장이다. 혹자는 퍼트남의 쌍둥이 지구논변이 보여주고 있는 것은 단어가 가리키는 외연을 결정하기 위하여 외부 대상의 본질을 고려해야 한다는 주장이지, 그 자체로서 외부세계가 우리의 심적 상태를 규정한다는 외재주의 논제인 것은 아니라고 반박할 수도 있을 것이다. 과연 그런가? 우리가 퍼트남의 쌍둥이 지구논변을 우리의 심적 표상 혹은 상태를 구분하기 위한 혹은 개별화하기 위한 동일성의 조건에 관한 것으로 이해한다면, 이는 전형적인 외재주의 입장이 된다. 가령 내가 '물'이라는 개념을 통하여 가지고 있는 표상은 쌍둥이 지구의 내가 가지고 있는 '물'이라는 표상과 같은 것인가 다른 것인가? 퍼트남의 쌍둥이 지구 논변의 핵심은 나와 쌍둥이 지구의 내가 가지고 있는 '물'에 대한 심적 표상은 그 내재적인 성질에 있어서는 동일하지만, 엄연히 서로 구분되는 다른 표상임을 주장한다. 물론 이때의 구분은 단순한 수적인 동일성에서의 구분을 이야기하고 있는 것이 아니다. 그러면 이때 내가 가지고 있는 표상과 쌍둥이 내가 가지고 있는 표상을 서로 다른 것으로 만들어 주는 것은 무엇인가? 물론 그것은 그 표상들의 내용의 차이이며, 그 내용의 차이는 바로 환경의 기여를 통하여 설명된다. 여기서 우리는 하나의 심리적 상태를 어떤 심리적 상태이게끔 만들어주는 하나의 구성적인 원리로서 외부적 환경이 이미 개입하고 있음을 관찰할 수가 있다.

어떤 의미에서 늑지인간이 지향적 태도를 갖느냐 갖지 않느냐의 문제는, 쌍둥이 지구인간이 가지고 있는 '물'에 대한 표상적 내용이 내가 가지고 있는 '물'에 대한 표상적 내용과 동일한 것인지 아닌지의 문제의 연

장선상에 놓여있다고 할 수 있다. 물론 쌍둥이 지구 논변과 높지인간의 논변의 경우에 중요한 한 가지 차이점이 있다. 쌍둥이 지구 논변에서 외부적인 환경은 일차적으로는 공간적인 외부 환경으로 제한된다. 하지만 높지인간과 관련한 목적론적 의미론은, 공간적인 의미에서의 외부적인 환경뿐만 아니라, 유기체와 관련한 역사적 사실을 이러한 외부적인 환경의 일부로 포함시킴으로써 더욱 강력한 형태의 외재주의를 채택하고 있다. 그런데 사실 이러한 차이는 표피적인 것일 수도 있다. 쌍둥이 지구 논변이 직접적으로 문제 삼는 내용은 자연종 명사와 같은 고정지시어들의 외연(의미) 결정이며, 쌍둥이 지구 논변 그 자체는 그렇게 결정된 의미들이 갖게 될 규범적 성격에 대해서는 아무런 언급도 하지 않는다. 만일 우리가 쌍둥이 지구논변 자체에서 이러한 의미의 규범적 성격을 문제 삼고자 했다면, 언어 공동체에서 통용되는 모종의 “지칭적 의도”와 같은 것이 언급되어야만 하고, 이러한 지칭적 의도를 비지향적, 비의미론적 개념으로 환원하는 길은 모종의 시간적, 역사적 고려를 통해야 할 것이라는 것이 필자의 생각이다. 필자의 이런 생각이 정당화 될 수 있다면, 쌍둥이 지구논변에서 시간적 역사성의 문제는 요구되지 않는 것이 아니라 단지 잠복된 채로 숨겨져 있을 뿐이다. 어찌 되었건, 현재 우리가 관심을 가지고 있는 외재주의는 공간적인 차원의 외부적인 환경뿐만 아니라 시간적인 차원의 외부적인 환경도 고려해야 한다는 아주 강력한 형태의 외재주의이다.

그렇다면 높지인간에 대한 목적론자들의 결론은 일차적으로 다음과 같이 변호될 수 있을 것이다. 가령 믿음과 같은 지향적 상태는 그 내용(content)을 통하여 동일성의 조건이 규정된다. 하나의 믿음을 그러한 믿음으로, 혹은 하나의 믿음과 다른 하나의 믿음을 동일한 믿음 상태로 만들어 주는 것은 바로 그 믿음의 표상적 내용이다. 그런데, 내용과 관련된 의미론적(semantical) 단계에서 지향적 상태를 구획 짓는 원리로서 우리가 가지고 있는 가장 유망한 이론은, 퍼트남이 쌍둥이 지구 논변을 통해 보여주었던 것과 같은 넓은 내용(wide content) 즉 외재주의적인 이론

이다. 만일 넓은 내용에 관한 쌍둥이 지구 논변을 받아들일 수 있다면, 이러한 입장에서부터 '높지 인간에게 나와 동일한 지향적 태도를 귀속시킬 수 없다'는 주장까지 가는 거리는 사실상 그렇게 멀지 않은 것일 수도 있다. 물론 이 주장에는 단서가 따른다. 앞서도 언급되었듯이, 우리의 목적론은 상당히 강력한 형태의 외재주의이며, 이러한 외재주의가 옹호되기 위해서는 선택의 역사라는 시간성을 요구하지 않는 온건한 형태의 외재주의와 차별화가 이루어져야만 하기 때문이다.

내용을 갖는 지향적 상태를 구획 짓는 원리의 문제에 있어서, 내재주의 입장에서 시작하여, 온건한 외재주의를 거쳐서, 목적론과 같은 강한 외재주의로 가는 철학적 동기를 다음과 같이 풀어볼 수 있을 것이다. 먼저 지향적 상태를 구획 짓는 것으로 내재주의적인 입장에서 채택할 수 있는 원리란 물리적인 인과적, 계산적 기능 혹은 성향적 속성과 같은 것들이다. 그러나 지향적 상태가 내용을 통하여 개별화 되는 한, 동일한 내용을 갖는 지향적 상태는 복수의 물리적, 기능적 상태에 의하여 실현될 수 있다. 우리는 이미 기능주의가 복수실현 가능성이란 논제를 통하여 전통적인 심신동일론을 논박하는 것을 목격하였다. 그런데 동일론을 대체한 기능주의란 일종의 내재적인 계산적 기계적 기능주의이다. 그러나 많은 사람들이 지적하고 있듯이 하나의 지향적 상태는, 이것을 실현하는 계산적 기능의 상태라는 측면에서도 또한 대단히 유연한(plastic) 것이다.<sup>8)</sup> 즉 동일한 내용을 갖는 하나의 지향적 상태는 복수의 계산적 기능 상태에 의하여 실현 가능하다. 만일 동일론이 지향적 태도의 구획에 대한 적절한 이론이 아니라면, 동일한 이유에 의해서 전통적인 의미에서의 기능주의적 이론도 적절한 이론이 아니다.

지향적 태도, 혹은 단순한 물리적 움직임과 구분되는 바로서의 지향적 행동이 단순히 종 한정적인(species specific) 개념이 아니라 여러 종에 걸쳐서 귀속될 수 있는 개념이라는 것을 고려하면 이점은 더욱 분명해진

---

8) 가령 Putnam (1988), ch. 5를 보라.

다. 만일 원숭이가 표현하는 '두려움'과 개가 표현하는 '두려움'을 함께 묶어줄 수 있는 구획의 원리라는 것이 있다면 그것은 도대체 어떠한 모습을 띄고 있겠는가? 이에 대한 내재주의자의 답은 대단히 웅색할 수밖에 없다. 지향적 설명의 대상이 되는 행위의 경우도 마찬가지이다. 드레츠키가 적절히 지적하고 있듯이, 지향적 설명의 대상이 되는 것은 행위 (behavior)이지 단순한 신체적인 움직임이 아니다.<sup>9)</sup> 이 말은 지향적 설명의 대상이 되는 행위가 단순히 외향적인 물리적 기계적 움직임만으로는 기술될 수 없다는 것이다. 가령 어떤 종류의 개똥벌레는 빛을 발함으로써 배우자를 유혹하지만 다른 종류의 개똥벌레는 빛을 발함으로써 먹이를 유혹한다. 이들은 외향적으로는 유사하지만 전혀 다른 범주에 속하는 행위들이다. 그런데 여기서 우리는 이 두 종류의 개똥벌레가 그 내재적인 인과적, 기능적 속성에 있어서는 완전히 동일한 가능성을 상상할 수 있다. 내재주의자는 이 두 가지 행위가 서로 다른 범주에 속한다는 것을 적절히 설명할 수가 없다.

여기서 한 가지 가능한 제안이 환경과의 관계를 포함하는 온전한 외재주의적 기능주의이다. 이 입장에 의하면 '눈이 희다'를 믿는 것과 같은 지향적 태도를 개별화 하는 기능적 역할은, 단순히 믿음 주체의 두뇌 상태뿐만 아니라 환경의 측면을 포함하는 식으로 확장되어야 한다는 것이다. 여기서는 단순히 우리의 두뇌가 아니라, 유기체가 속하는 사회적 물리적 환경을 포함하는 전체가 지향적 상태의 기능적 역할을 규정하는 시스템이 된다. 지향적 상태의 내용은 이제 유기체의 두뇌상태와 외부세계의 대상 사이에 성립하는 모종의 기능적 인과관계를 통하여 분석될 것이다. 하나의 심리적 상태를 그러한 심리적 상태로 만드는 것은 단순한 좁은 의미에서의 계산 기능적 상태를 실현하고 있다는 것이 아니라, 환경과의 상호작용을 포함하는 넓은 의미의 계산적, 인과적 기능이라는 것이다. 그런데, 설명 기능을 이런 넓은 방식으로 이해한다고 하더라도 한 가

---

9) Dretske (1988), ch. 1.

지 변하지 않는 점이 있다. 그것은 이런 넓은 기능의 개념조차도 단지 공간간적인 인과적, 성향적 분석을 통해 도출되는 것이며, 그런 한에 있어서 여전히 기술적인 성격을 띠고 있다는 점이다. 우리는 앞서 지향적 내용이란 마땅히 그러해야만 하는(ought to, supposed to) 바로서의 규범적 성질을 띠고 있다고 가정했다. 그런데, 넓은 기능의 분석이 그 기본적인 성격상 기술적일 수밖에 없다면, 어떻게 이로부터 지향적 내용의 규범적 성격을 확보할 수 있을 것인가?

역사의 기여를 포함하는 목적론적 개념으로서의 기능 개념의 도입이 이루어지는 것은 바로 이 지점에서이다. 발생론적 기능개념을 주장하는 사람들의 동기는, 역사적 고려가 없는 단순한 넓은 의미의 기능개념만으로는 지향적 태도가 갖는 규범적 성격을 확보하기 어렵다는 것이다. 어떤 현상이 규범적이라는 것은 강한 목적성을 전제로 한다. 목적론자들은 기술적인 성격의 넓은 기능개념만으로는 이런 목적성을 끌어 낼 수 없다고 생각한다. 이들에 따르면, 지향적 태도에 어떤 목적성을 부여할 수 있는 길은, 오직 이들을 진화를 통한 선택과 같은 역사적 관점에서 접근할 때이다. 우리가 지향적 태도를 통하여 행하는 지향적 행위는, 단순히 어떤 기술적인 의미에서의 인과적, 기능적 상호작용만을 하고 있는 것이 아니라, 가령 환경에 대한 정보를 습득하고 환경의 변화에 대처함으로써 살아남으려고 하는 모종의 목적성을 전제로 한 행위라는 것이다. 그리고 맥락적이거나 관심상대적이 아닌 보다 궁극적인 의미에서의 목적성을 확보하는 길은, 기능이라는 것을 단순히 공간간적 개념이 아니라, 통시간적인 개념으로 보는 것이다. 이제 이러한 입장에서, 어떤 지향적 상태의 수반적 기초는 단순히 그 상태가 귀속되는 시스템 내외부의 물리적이거나 계산적인 상태가 아니라, 그 시스템을 포함하는 통시간적인 생태학적(ecological) 공간이 된다.

한편으로, 목적론적인 의미론자들이 취하고 있는 지향적 태도에 대한 발생론적 접근 방식은 지향성의 개별화와 관련한 외재주의적인 입장에 대한 보다 폭 넓은 동기를 제공하고 있는 것처럼 보인다. 발생론적 관점

에서 볼 때, 믿음이나 욕구와 같은 인지 기제들은 먹이를 찾거나, 약탈자를 회피하는 것과 같은 환경적 압력을 극복하기 위해 선택된 적응(adaptation)의 산물이다. 이러한 지향적 상태들이 어떠한 특징을 가지고 있는 것은 결국 이들 상태의 주체가 어떤 특정한 환경에 처해졌었기 때문이며, 만일 그 주체가 다른 환경적 조건에 처해졌다면 다른 특징을 갖는 다른 상태를 발전시키지 않았겠는가? 이런 점에서, 이들 지향적 상태들이 갖는 본성은 본질적으로 관계적인 것이다. 만일 지향적 상태의 본성이 관계적인 것이라면 심성내용의 외재성 혹은 지향적 태도의 구획에 있어서 외재주의는 당연한 귀결이다.<sup>10)</sup>

3

마지막으로 이 절에서는 일련의 사유실험을 통하여, 넓은 기능의 온전한 외재주의로는 지향적 상태가 갖는 규범적 성격을 해명하기 어려우며, 이를 위해서는 기능에 대한 발생론적 이해가 필수적으로 요구된다는 목적론자의 주장을 보다 적극적으로 옹호해 보자. 만일 지향적 태도가 그 구성적 본질상 규범적인 것이며, 이를 확보하는 유일한 길이 기능에 대한 발생론적 이해뿐이라면, 목적론에서 말하는 발생적 기능개념은 최상의 설명이라는 추리의 일환으로서 정당화 될 수 있다. 그런 한에 있어서, 발생론적 기능개념은 지향적 태도의 본질적 구성요건으로 편입될 것이

---

10) 맥킨은 이런 맥락에서 발생론적인 목적론이 쌍둥이 지구논변과 관련된 지표성(indexicality)에 대하여 매우 흥미로운 설명을 제공하고 있음을 지적한다. 목이 말라서 나는 물을 마시고 싶어 한다. 그런데, 내 위나 간의 기능은 H<sub>2</sub>O를 처리하는 것이다. 이들 기관들이 갖는 특질들은 내가 처해있는 이 지구라는 환경 속에서 발견된 어떤 특정한 물질을 이용하기 위하여 진화된 것이기 때문이다. 그런 점에서, 물에 대한 나의 표상은 지구상의 바로 그 물질을 가리킨다. 한편으로 쌍둥이 지구의 나에게 속한 위나 간의 한번도 H<sub>2</sub>O를 접한 적이 없으며 쌍둥이 지구에서 발견되는 특정한 환경적 조건 즉 XYZ를 이용하기 위해 진화된 적응물이다. 그런 점에서, 쌍둥이 지구인이 갖는 물에 대한 표상은 XYZ를 가리켜야만 한다. (McGinn, (1989), pp.157-158.)

고, 늙지인간에게 아무런 지향적 태도도 귀속 시킬 수 없다는 목적론자의 주장도 정당화될 것이다.

먼저 다음과 같은 예를 상상해보자. 바닷바람이 불어서 아니면 개미떼가 지나간 흔적에 의하여 정말 우연적으로 백사장에 ‘퍼트남’이라는 모양의 형체가 생겨났다고 가정해 보자.<sup>11)</sup> 이 형체는 그 자체로서 퍼트남이라는 사람을 가리키는 이름인가, 아니면 그냥 아무런 의미가 없는 물리적 형상일 뿐인가? 만일에 누군가가 의도적으로 모래사장에 ‘퍼트남’이라는 글자를 써놓았다면 이는 분명히 퍼트남이라는 사람을 가리키는 이름일 것이다. 그런데, 개미떼가 우연히 만들어 놓은 ‘퍼트남’이란 형상은 그 물리적, 인과적 성질에 있어서는 누군가가 써놓은 ‘퍼트남’이란 글자와 동일함에도 불구하고 퍼트남을 가리키는 이름으로 간주될 수 없다. 이것은 아무런 의미론적 성질을 갖지 못하는 물리적 형상일 뿐이다. 이 두 가지 경우의 차이는 발생론적인 차이이다.

만일 이 ‘퍼트남’이라는 형상에 의미를 부여할 수 없다는 주장에 동의할 수 있다면 다음의 경우는 어떠한가? 어느 날 바람이 불고 번개가 치더니, 나무에 벼락이 떨어졌다. 어떤 우주적인 우연에 의하여 주위에 널려있던 분자가 합성되어 반인반마의 모습을 한 켄타우로스 같은 존재가 생겨났다고 가정하자. 그런데, 이 존재는 가령 ‘물’과 같은 인간이 사용하는 말과 비슷한 몇 마디의 소리도 낼 수 있다. 덧붙여 이 존재는 물을 보면 ‘물’이라는 소리를 내기도 하고, 그런 소리를 낼 때마다 우리가 물을 가져다주면 그 물을 ‘마시기’도 한다. 하지만 이 존재는 소위 퍼트남이나 크립키적인 의미이론에서 말하는 어떠한 언어적 공동체에 속하지도 않으며, 우리가 사용하는 ‘물’이라는 말과 아무런 인과적 연쇄관계도 맺고 있지 않다. 이 존재는 말 그대로 갑자기 우연적으로 생겨나서 ‘물’이라는 소리를 내고 있는 것이다. 그렇다면 이 존재가 내는 ‘물’이라는 소리는 표상적 내용을 갖는 언어적인 표상인가, 아니면 아무런 의미도 갖

---

11) Putnam (1981), p.1 참조.

지 않는 소음에 불과한 것일까? 만일 이것이 언어적 표상이라면, 그것은 우리가 H<sub>2</sub>O를 가리키기 위하여 사용하는 ‘물’이라는 말과 동일한(혹은 유사한) 표상적 내용을 갖는가, 아니면 그것과는 구분되는 전혀 다른 의미를 가지고 있는가?

만일 이것이 우리가 사용하는 ‘물’이라는 표현과 비슷한 언어적인 표상이며, 이 존재는 그에 상응하는 모종의 지향적 상태를 가지고 있다고 가정해보자. 그러면 이것이 콜라를 보고 ‘물’이라는 소리를 내는 것은 올바른 언명인가, 아니면 거짓 믿음과 같은 오표상(misrepresentation)의 사례인가? 만일 이것이 오표상의 경우라면, 무엇이 이것을 오표상으로 만드는가? 먼저 지향적 내용이 역사성을 배제한 물리적 환경과의 인과적 상호작용에 의해 고정될 수 있다는 온건한 의미의 외재주의자의 입장을 고려해 보자. 이러한 입장은 이른바 긴팔 기능주의(long-arm functionalism)라고 부를 수 있는 입장으로서, 앞서도 언급하였듯이 기능적 역할 분석의 대상이 되는 시스템을 물리적 환경을 포함하는 방식으로 확장하려는 시도이다. 그런데 만일 이러한 온건한 외재주의자가 켄타우로스의 표상적 상태를 기능적 상태로 정의한다면, 이들이 호소할 수 있는 유망한 기능 개념은 일종의 커밍스적인 기능 개념이다.<sup>12)</sup>

커밍스는 어떤 상태 혹은 사물이 현재 가지고 있는 성질 혹은 능력을 통하여 기능개념을 분석한다. 우리가 보통 기능주의를 말할 때의 그 기능 개념은 이런 커밍스의 분석에 매우 가까운 것이다. 커밍스는 자신의 이론을 ‘기능적 분석(functional analysis)’의 이론이라고 부르는데, 이 이론에 따르면 어떤 대상 X에 모종의 기능 F를 귀속시키는 진술의 목적은, X가 속해 있는 어떤 복잡한 시스템의 능력 혹은 행위를 설명하기 위한 것이다. 가령 X가 속해 있는 시스템 S가 어떤 복잡한 능력 C를 갖는다고 하고, S가 어떻게 하여 그런 복잡한 능력 C를 갖게 되었는가를 설명하려고 한다고 해보자. 이때 이 복잡한 능력 C를, X를 포함한 S를 구성하고

---

12) Cummins (1975).

있는 부분들이 행하는 보다 단순한 행위 혹은 능력들로 분해하고, 이런 단순한 것들이 모여서 어떻게 복잡한 능력 C를 갖게 되는 가를 설명하려는 것이 바로 기능적 진술의 목적이다. 다시 말해서, X가 어떤 기능 F를 갖는다는 것은, X가 F라는 인과적 역할을 수행함으로써 X가 속한 시스템 S가 복잡한 능력 C를 가짐에 기여한다는 것을 의미한다. 그런데 여기서 커밍스의 기능이론이 역사적인 기능이론과 구분되는 중요한 한 가지 차이점은, 어떤 X가 커밍스적인 기능 F를 갖기 위해서는, F를 수행하기 위한 인과적 혹은 성향적 속성을 실제로 가지고 있어야 한다는 점이다. 다시 말해서, 커밍스의 기능개념이란 결국 성향적 분석이다. 만일 X가 가지고 있는 현재의 인과적, 성향적 속성이 F를 수행하기에 무력하다면, F라는 기능을 X에 귀속시킬 수는 없다.

커밍스 이론의 이러한 특징은, 기능에 대한 역사적 접근에 비추어 규범성의 해명이라는 측면에서 중대한 차이를 가져온다. 역사적 이론은 기능이 과거의 주어진 역사에 의해 결정된다는 것을 주장함으로써 우리의 설명적 관심으로부터 독립적인 목적성을 도입하고, 규준과 실제의 괴리를 확보할 수 있었다. 하지만 커밍스의 이론에서는 규준과 실제의 괴리에 대한 설명이 그렇게 용이하지 않다. 가령 병이 든 심장의 예를 생각해 보자. 병이 든 심장의 경우 이것은 이미 혈액을 순환시킬 수 있는 인과적 성질이나 성향을 갖지 못한다. 그러므로 우리는 이 병든 심장의 기능이 혈액을 순환시키는 것이라고 말할 수 없다. 동시에 우리는 병든 심장이 그 정상적 기능을 수행하지 못하는 일탈적인 것이라고 이야기할 수도 없다. 대신에 이 병든 심장은 여전히 어떤 복잡한 시스템의 한 부분일 수 있다. 가령 극단적으로 어떤 유기체의 죽음이라는 복잡한 과정에 기여하는 것이 그것의 기능일 수 있다. 커밍스의 분석에 있어서 분석의 대상이 되는 시스템은 꼭 어떤 유기체의 생존에 꼭 필요하다거나 하는 중요한 행위를 수행해야 하는 것은 아니다. 커밍스 이론의 목적은, 그것이 무엇이 되었든 간에 어떤 복잡한 시스템의 행위나 능력을 그 부분들의 단순한 행위나 능력을 통하여 설명하는 데에 있기 때문이다. 또한,

커밍스의 분석에 있어서 분석의 대상이 되는 시스템에 어떤 목적성을 부여할 수 있다면, 그 목적성은 기능적 분석을 통하여 우리가 설명하고자 하는 것이 무엇이나에 의존하는 일종의 설명 상대적인 목적성이다. 이는 디자이너의 의도에 근거하여 인공물에 목적성을 부여하는 것에 유사한, 일종의 파생적인 목적성이다. 그러나 우리의 목적이 파생적 지향성에 대비되는 본래적 지향성의 규명인 한에 있어서, 우리에게 필요한 것은 그런 설명 상대적 목적성이 아니라 모종의 본래적인 목적성이다. 이런 본래적인 지향적 태도가 가지는 규범성을 경험과학적인 자연적 성질을 통하여 해명하는 것이 자연주의자가 가지고 있는 중요한 화두 중의 하나라고 한다면, 이런 점에서 커밍스의 기능개념은 결코 매력적인 선택이 아니다.

그런데, 문제는 온건한 외재주의가 호소할 수 있는 기능개념이란 바로 이런 커밍스적인 기능개념이라는 것이다. 가령, 온건한 외재주의자는 켄타우로스가 내 뱀을 '물'이라는 발화와  $H_2O$  사이에 성립하는 모종의 물리적, 인과적, 기능적인 관계에 의하여 '물'이라는 발화 혹은 이와 관련된 지향적 태도의 내용이 결정된다고 주장할 것이다. 그런데, 이 켄타우로스가 콜라를 보고도 '물'이라고 내뱀었다고 가정해 보자. 이 때 온건한 외재주의자는 이것이 오표상의 사례임을 어떻게 확보할 수 있는가? 켄타우로스가 발화한 '물'이라는 언명과 콜라 사이에도 분명히 모종의 성향적, 기능적 관계가 존재할 것이다. 그렇다면, 이때 온건한 외재주의자가 져야 될 부담은, '물'과 물 사이에 성립하는 기능적 관계와, '물'과 콜라 사이에 성립하는 기능적 관계를 차별화시켜야 하고, 왜 전자의 기능적 관계는 마땅히 그러해야 하는 바로서의 규범적 지위를 누리게 되지만 후자의 기능적 관계는 그렇지 못한 지를, 두 기능적 관계 사이에 존재하는 차이를 통해서 보여주어야만 한다. 그런데, 그러기 위해서는 전자의 기능적 관계가 그 켄타우로스가 '의도'한 것이라는 취지의 강한 목적성이 도입되어야 한다. 문제는 그러한 의도나 목적성을 비록 넓은 의미의 기능개념을 끌어들이는 하더라도 성향적인 기능적 분석만으로는 부여할 방법이 없다는 것이다. 극단적인 한 예로, 우리는 이 켄타우로스가 이 지구상에서의 성

공적인 생존을 지향하고 있다고 생각할 아무런 이유도 없다. 비근한 예로 우리 사람들 중에는 자살 충동을 가지고 있는 유전적 인자를 가지고 있는 사람도 있다고 하지 않는가? 켄타우로스에 어떤 목적성을 부여하지 않는 것은, 우리가 돌맹이를 하늘로 던졌을 때, 법칙적(nomological) 의미와는 구분되는 규범적인 의미로, 그 돌맹이가 어떠한 궤적을 그리며 날아가야만(should) 된다고 말하지 않는 것과 유사한 이유이다. 비록 켄타우로스가 모종의 인과적 기능적 상태를 실현하고 있는 것으로 기술될 수는 있지만, 그것이 궁극적인 차원에서 어떤 모종의 상태를 지향하고 있다고 우리는 생각해야 할 이유는 없다. 만일 그렇다면, 콜라에 대한 ‘물’이라는 켄타우로스의 언명을 오표상으로 만들거란 요원해 보인다. 크립키의 논점을 차용하자면, 이때 온건한 외재주의자가 처한 상황은, 마치 성향적 분석을 통하여 커하기(quus)에 대해서 더하기가 갖는 우선성을 확보하려는 성향론자의 처지와 같다.

온건한 외재주의자가 처한 곤궁을 더 드러내기 위하여, 사유실험을 조금만 더 진행시켜 보자. 이제 우리의 켄타우로스가 높지에서 생겨난 지 얼마 지나지 않아서 쌍둥이 지구로 이동하였다고 가정해보자. 여기서 일단 잠정적으로 이 존재가 지구에서 머문 기간이 너무나 짧아서 H<sub>2</sub>O와 충분한 인과적 상호작용을 갖지도 못했고, 지구상의 언어공동체에 속하지도 않기에 우리로부터의 어떠한 형태의 의미 빌립(reference borrowing)도 가능하지 않았다는 것을 가정해보자. 그러면 이제 쌍둥이 지구상에서 이 존재가 내뱉는 ‘물’이라는 발화는 H<sub>2</sub>O를 가리키는가 아니면 XYZ를 가리키게 되는가? 혹자는 여기서 켄타우로스가 지구에서 먼저 태어났으므로 이것이 발화하는 ‘물’이라는 언명은 H<sub>2</sub>O를 가리키고, 그러므로 쌍둥이 지구에서 발화하는 ‘물’이라는 언명은 XYZ에 잘못 적용된 거짓언명이라고 주장할 수 있을 지도 모르겠다. 하지만 이러한 답변은, 앞 절에서 잠시 언급되었던, 쌍둥이 지구논변에 숨겨져 있는 지칭적 의도란 것을 무시한 발상이다. 우리가 발화하는 ‘물’이라는 언명이 H<sub>2</sub>O를 가리키게 되는 것은 단순히 지구상에 있는 액체가 H<sub>2</sub>O라는 분자구조를 가지고 있기

때문만이 아니라, 우리가 '물'이라는 언명을 통하여 그 액체와 동일한 분자구조를 가지고 있는 대상을 가리키고자 하는 지칭적 의도가 있었기 때문이다.

그러면 문제는 어떠한 방식으로 켈타우로스에게 그러한 지칭적 의도를 부여할 수 있는가? 켈타우로스의 경우 어떤 명시적인 명명식(baptism)이 결부되어 있지는 않다. 그렇다면, 이 존재가 H<sub>2</sub>O와 갖게 되는 최초의 인과적 상호작용이 '물'이라는 발화의 내용을 고정하게 되는가? 그것이 무엇이 되었든 간에, 켈타우로스의 어떤 언명과 대상 사이에 성립하는 단순한 최초의 인과관계를 통하여, 앞으로 그 언명이 따라야 할 규칙, 규범으로서의 의미가 결정된다는 것은, 마치 켈타우로스의 언명에 어떤 마법적인 능력을 부여하는 것과 같아 보인다. 보다 유망한 전략은 아마도 충분한 정도의 인과적 상호작용에 호소하는 방식일 것이다. 그런데, 도대체 몇 회 정도의 상호작용이 있어야 충분한 상호작용으로 간주될 수 있는가? 물론 나는 그 대답을 알지 못한다. 하지만 내가 그 답을 모르고 있는 것은 내가 꼭 알아야만 하는 어떤 인과적, 기능적 사실에 무지하여서가 아니라, 정답이 없기 때문일 것이다. 그것이 몇 회가 되었건 간에, 그런 횟수를 통하여 오표상을 결정짓는 규범으로서의 의미를 결정하는 것은 임시변통(ad hoc)적인 것으로 보인다. 선택의 과정이 개입되지 않은 단순한 횟수의 축적이, 마땅히 표상하도록 되어있는(supposed to represent) 바로서의 규범적인 성격을 확보해 주지는 않는다.

여기서 혹자는 높지 인간은 우리가 속한 언어공동체에 편입되면서, 모종의 의미 빌립 혹은 공유가 일어날 수 있기에, 켈타우로스의 예와는 본질적으로 다르지 않은가 하고 의문을 제기할 수 있다. 과연 그런가? 일단 이들이 태어난 바로 그 순간만의 시점을 생각해보자. 이들 사이에 규범적인 의미에서의 지향적 태도의 귀속과 관련한 어떤 본질적인 차이가 있는가? 위에서 잠시 언급되었듯이 무엇에 대한 징표를 갖는다는 것과 실제로 그러하다는 것은 다른 문제이다. 그런데, 높지 인간의 경우 우리는 그것이 가지고 있는 우리와의 유사성 때문에 모종의 편견을 가질 수

도 있다. 굳이 켄타우로스의 예를 끌어 들인 것은 우리나라 혹은 지구상에 생존하는 다른 생물과의 유사성을 배제함으로써 그러한 편견의 개입을 최대한 억제하기 위한 것이었다. 만일 우리가 그러한 편견을 배제한다면, 최소한 늑지인간과 켄타우로스 사이에는 어떠한 본질적인 차이도 없어 보인다. 만일 태어난 바로 그 시점에 켄타우로스에게 어떤 목적성도 부여할 수 없으며, 따라서 규범적 평가의 적용을 받은 지향적 상태를 귀속시킬 수 없다면, 태어난 바로 그 시점의 늑지인간의 경우도 마찬가지일 것이다. 여기서 이들 사이에 어떤 본질적인 차이가 있다고 주장하는 것은 이미 논점을 선제하는 선결문제의 오류를 저지르는 것처럼 보인다.

늑지인간이 우리의 언어공동체에 편입됨으로써 모종의 의미 빌림이 일어난다는 주장은 어떠한가? 도대체 그러한 일은 어떻게 일어나는가? 어떤 상호약정이라도 체결되는 것일까? 목적론자는 아마도 늑지인간이 우리 언어공동체의 일원이라는 것을 부정할 것이다. 우리의 언어 공동체는 일종의 규범적공동체이다. 만일 늑지인간이 우리 언어공동체의 일원이라고 생각할 수 있다면, 이는 늑지인간의 인과적, 성향적 속성이 우리 언어공동체가 채택하고 있는 언어적 규범에 부합하는 방식으로 보이기 때문일 것이다. 그런데, 목적론자들이 굳이 발생론적 기능을 끌어들이려고 하는 이유는, 이러한 인과적, 성향적 속성의 부합만으로는 늑지인간의 지향적 상태나 행위를 규범적인 것으로 만들기 충분하지 않다고 생각하기 때문이다. 규칙따르기와 관련하여 흔히 드는 비유적인 예로 체스게임을 생각해 보자. 체스 게임은 일종의 규칙들로 이루어져 있으며, 그 경기에 참가하는 자들은 이 게임의 규칙을 따르고자 한다. 그런데, 체스의 규칙을 전혀 이해하고 있지 못하지만, 체스의 규칙에 일치하여 말을 움직이는 사람이 있다고 가정해보자. 옆에서 보았을 때, 이 사람은 체스의 규칙을 이해하고 있는 듯이 보이고 또 체스경기를 하고 있는 것처럼 보이지만, 이 사람은 게임을 하고 있는 것이 아니라 단순히 자신의 성향에 따라서 말을 움직이고 있을 뿐이다. 이때 우리는 이 사람이 그 정당한 의미에 있어서 체스 경기를 하고 있다고 말할 수 없다. 늑지인간의 인과적, 성향

적 속성이 규범적 평가의 적용을 받지 않는다는 목적론자들의 주장에는 바로 이러한 사고가 깔려 있다. 높지인간의 지향적 상태에 대하여 어떤 목적성을 부여할 수 없는 한, 그것이 보이고 있는 상태나 행위는 단지 우연적으로 규칙에 부합하는 행위일 뿐, 규칙을 따르고 있는 상태나 행위가 아니라는 것이다. 그런 한에 있어서 높지인간은 우리 언어 공동체의 일원이 아니다. 나는 목적론자들의 이러한 생각이 보고시안이 정리하고 있는 규칙따르기에 대한 크립키의 입장과 크게 다르지 않다고 생각한다. 보고시안은 크립키의 견해를 다음과 같이 평가한다.<sup>13)</sup>

크립키는, 어떤 표현의 외연을 정확히 포착해내는 적절하게 선택된 성향이 설령 있다고 하더라도, 그러한 성향은 의미에 대한 사실과는 여전히 동일시 될 수 없다고 생각하는 것 같다. 성향은 기술적인 개념인 반면에 의미가 기술적인 개념이 아니라는 것은 여전히 변함없이 참으로 남아 있기 때문이다. 다시 말해서, 크립키에 따르면, 의미술어와 동변적인(covaryed with) 성향 술어가 있다고 하더라도, 각각에 관련된 사실은 서로 다른 종류의 사실이기에 서로 동일시 될 수 없다는 것이다.

만일 크립키의 이런 주장을 받아들일 수 있다면, 높지인간이 우리의 언어공동체의 일원이라고 주장하기 위한 가장 손쉬운 길은 우리의 지향적 상태나 내용이 규범적이라는 것을 부정하는 것일 것이다. 그러나 우리는 이미 이 전체를 받아들이기로 결정하였다. 그렇다면, 온전한 외재주의자가 높지인간이 우리 언어공동체의 일원이라는 것을 주장하기 위하여 짊어져야 할 부담은, 단순히 규칙에 부합하는 인과적 성향이나 속성을 갖는다는 것으로부터 어떻게 온전한 의미에서의 규칙따르기가 도출될 수 있는가를 보여주는 일이다. 목적론자들은 바로 이러한 시도가 성공할 수 없다는 것을 주장하고 있는 셈이다.

우리는 지금까지 일련의 사유실험과 또 그에 따른 논의를 통하여 높

13) Boghossian (1989), p.532.

지인간과 관련된 목적론자들의 핵심 주장이 무엇인지를 살펴보았다. 그리고 그 논의의 결론은, 어떤 존재에 대한 적절한 선택의 역사가 없이는, 그 존재의 지향적 상태가 무엇을 나타내도록 되어있는지의 목적성에 대하여 원칙적인 답을 해줄 수 없다는 것이다. 만일 이러한 목적성으로서의 규범성이 지향적 태도를 구성하는 본질적인 요건이라면, 이러한 목적성의 요건을 만족시키지 못하는 늑지인간에게는 아무런 지향적 상태도 귀속시킬 수 없다. 물론 이러한 가상의 예들은 양쪽의 날을 가진 칼 같은 것이라고 할 수 있다. 가령 이러한 예들은 데넷과 같은 (아마도 콰인을 포함하여) 사람들이 주장하듯이 지향성이나 의미라는 것은 실재하는 것이 아니라 우리가 어떤 시스템의 움직임을 예측하고 설명하기 위한 도구적인 성격만을 지닐 뿐이라는 결론을 유도하기 위하여 사용될 수도 있다. 하지만 우리가 지향성이나 의미와 관련한 모종의 실재론을 전제한다면, 역사성이 결여된 단순히 그 존재의 인과적, 기능적, 행동적 성향만으로 오표상을 논할 수 있는 규범적인 차원의 의미를 부여할 수 있을 것 같지는 않다. 켄타우로스나 늑지인간 같은 존재가 내뱉는 ‘물’이란 소리에 국한하여 말한다면, 목적론자들은 규범적 의미를 결정지어 줄 수 있는 아무런 의미론적 사실도 없다는 크립키의 회의주의자의 주장에 동의할 것이다.

마지막으로 늑지인간이 아무런 지향적 태도를 가지지 않는다고 말할 때, 목적론자들이 부정하는 것이 정확히 무엇인지를 정리하면서 이 논문을 마무리 하자. 앞서도 지적하였듯이 이들이 노리고 있는 것은, 인간이나 다른 유기체의 행위를 설명하기 위한 지향적 심리학에서 채택하고 있는 믿음이나 욕구와 같은 상태들에 대한 일종의 유망한 이론적 정의를 제시하는 것이다. 이들 상태들은 그것들이 갖는 내용을 통하여 구체되는 상태들이며, 본질적으로 이들이 표상하는 내용에 근거한 규범적인 평가의 적용을 받는 상태들이다. 규범적 성격을 지닌다는 것은 언제나, 어떤 것이 마땅히 그러해야 한다거나 그러할 것으로 기대되는 일종의 목적성이 전제되는 개념이다. 이들의 주장은 이러한 목적성을 해명하기 위해서

는 역사적 시간성이라는 외부적인 사실의 개입이 필수적이라는 것이다. 높지인간에게서 부정되고 있는 것은 이러한 규범적 평가의 적용을 받는 태도로서의 지향적 상태들이다. 목적론자들도 높지인간의 내부에 우리와 유사한 모종의 물리적인 인과적 과정이 진행되고 있으며, 그러한 과정을 통하여 어떤 외부적인 물리적 움직임이 촉발되고 있음을 부정하지 않는다. 하지만 그러한 인과적 과정이 아무리 우리와 유사하다고 하더라도, 그런 인과적 과정에 대하여 아무런 목적성도 부여할 수 없다. 그 결과 우리는 가령 그것이 참이거나 충족되었다는 식의 규범적 평가를 내릴 수 없다는 것이다. 이런 식으로 이해된 높지인간에 대한 목적론자들의 주장은 대단히 온건한 주장처럼 보인다.

참 고 문 헌

- 신상규 (2003a), “의미론적 규범의 자연화에 대한 목적론적 전략”, 철학과 현실, 여름.
- \_\_\_\_\_ (2003b), “믿음 내용의 고정에 대한 진화론적 접근”, 분석철학회 여름세미나.
- Boghossian, P. (1989), “The Rule-Following Considerations”, *Mind* 98: 507-549.
- Buller, D. J. (1999), *Function, Selection and Design*, Albany: SUNY Press.
- Burge, T. (1986), “Individualism and Psychology”, *Philosophical Review* 95: 3-46.
- Cummins, R. (1975), “Functional Analysis”, *Journal of Philosophy* 72: 741-765.  
Reprinted in Buller (1999).
- Davidson, D. (1987), “Knowing One's Own Mind”. Reprinted in Davidson (2001), *Subjective, Intersubjective, Objective*, Oxford: Clarendon Press.
- Dennett, D. C. (1987), *The Intentional Stance*, Cambridge: MIT Press.
- Dretske, F. I. (1986), “Misrepresentation” in Stich and Warfield (1994).
- \_\_\_\_\_ (1988), *Explaining Behavior*, Cambridge: MIT Press.
- \_\_\_\_\_ (1997), *Naturalizing the Mind*, Cambridge: MIT Press.
- \_\_\_\_\_ (2000), *Perception, Knowledge, and Belief: Selected Essays*, Cambridge: Cambridge University Press.
- Kripke, S. A. (1982), *Wittgenstein on Rules and Private Language*, Oxford: Basil Blackwell.
- McGinn, C. (1989), *Mental Content*, Oxford: Basil Blackwell.
- Millikan, R. G. (1984), *Language, Thought, and other Biological Categories*, Cambridge: MIT Press.
- \_\_\_\_\_ (1993), *White Queen Psychology and Other Essays for Alice*, Cambridge: MIT Press.
- Putnam, H. (1975), “The Meaning of “Meaning”” in *Mind, Language and*

*Reality.*

\_\_\_\_\_ (1981), *Reason, Truth and History*, Cambridge: Cambridge University Press.

\_\_\_\_\_ (1988), *Representation and Reality*, Cambridge: MIT Press.

Stitch, S. P. and Warfield, T. A. eds. (1994), *Mental Representation: A Reader*, Cambridge: Blackwell.

## Swampman and Intentionality

Sangkyu Shin

According to the teleosemantic theory of representational content, representation states are proper functional states. But, having a proper function is essentially a matter of evolutionary history. This leads us to a conclusion that a 'swampman' that suddenly arises in a swamp by improbable cosmic incidence but otherwise a perfect physical duplicates of a normal person, doesn't have any intentional states since it had lacked any history. Some people think that this conclusion is very counter-intuitive. I approach this objection with the distinction between externalism and internalism, and examine what sort of background theory makes the swampman case look convincing. I argue that, if we assume that 'intentional states' are normative phenomena, the seemingly counter-intuitive conclusion for the swampman held by teleosemanticists can be defended as an inference to the best explanation of this normative character.

**[Subject]** Philosophy of Mind, Philosophy of Language

**[Key Words]** intentionality, naturalization of intentionality, normativity of intentionality, teleosemantics, externalism, swampman