# 의미론적 역설과 진리 개념*

이 종 권 (중앙대)

**주제분류** 논리 철학

**주 요 어** 의미론적 역설, 거짓말쟁이의 역설, 진리론, 타르스키, 크립키

**요 약 문**

대부분의 철학자들은 거짓말쟁이의 역설이 진리 개념을 새로 정의할 것을 요구하는 것으로 생각하고 있다. 진리 개념에 대한 탐구를 시도한 대표적인 인물인 타르스키와 크립키도 그러한 철학자들에 속하는데 그들은 '진리'의 의미를 새로 규정한 다음 그러한 규정에 비추어 거짓말쟁이의 역설과 같은 것이 더 이상 야기되지 않음을 보이려 했다. 이 글에서는 그들의 진리 정의를 비교해서 검토하고 그들의 진리 개념이 어떤 난점을 안고 있는가를 살피고 있다,

* 이 논문은 2003년도 한국 학술 진흥 재단의 지원에 의하여 연구되었음. (KRF-2003-074-AS0018)

## 1. 타르스키의 진리 정의와 거짓말쟁이의 역설

의미론적 역설이란 의미론적 개념과 관련된 역설을 말한다. 20세기 시작을 전후하여 발견된 의미론적 역설로서는 거짓말쟁이의 역설(The Liar Paradox), 그밖에 리샤르의 역설(Richard Paradox), 베리의 역설(Berry Paradox) 등이 있는 데, 그 가운데 대표적인 것은 거짓말쟁이의 역설로서 의미론적인 역설에 관한 대부분의 논의는 이 역설에 집중되어 있다. 거짓말쟁이의 역설은 진리 개념과 관련된 역설이다. 여기서 진리는 세계와 문장과의 관계 덕분에 문장이 지니게 되는 속성으로 간주된다. 그러한 속성을 가리키는 일항의 술어를 $T$라고 하자. 그 술어가 적용될 자리에는 문장을 가리키는 이름이 올 것이다. 문장 $\phi$를 나타내는 이름을 '$\phi$'라고 하면 $T$('$\phi$')는 "문장 $\phi$가 참이다."라는 것을 의미한다.

술어 $T$의 의미 내지는 그 개념을 파악하기 위해서는 그 술어의 외연을 정의해야 할 필요가 있다. 진리 술어의 외연을 정의함으로써 그 개념을 파악하려 처음으로 시도한 것은 타르스키(A. Tarski)였다. 그는 진리 개념의 정의가 만족할만한 것이 되기 위해서는 진리 술어 $T$가 다음을 만족해야 한다는 생각에 대한 검토부터 시작하고 있다.

**A**. $T(\text{‘}\phi\text{’}) \leftrightarrow \phi$

실질적인 적합성의 조건에 의하면 진리 술어 $T$의 외연은 참인 문장을 가리키는 이름으로 구성된다. 그런데 A를 만족하는 진리 술어는 지기지시적인(self-referring) 문장과 관련하여 역설을 야기한다. 자기지시적인 문장이란 자기 자신에 대해 어떤 주장을 하는 문장이다. $\psi$('$\phi$')는 문장 $\phi$가 $\psi$임을 주장하는 문장이다. 그런데 문장 $\psi$('$\phi$')가 바로 $\phi$ 자신이라고 한다면, 즉 $\psi$('$\phi$')=$\phi$라면 다음이 성립할 것이다.

**S**. $\psi(\text{‘}\phi\text{’}) \leftrightarrow \phi$

A와 S로부터 다음이 귀결된다.

(1) $T(‘\phi’) \leftrightarrow \psi(‘\phi’)$

만일 ψ가 거짓 개념을 표현하는 술어라면, 즉 $F$라면 (1)은 다음과 같은 형태가 될 것이다.

(2) $T(‘\phi’) \leftrightarrow F(‘\phi’)$

(2)는 ϕ가 참일 경우 또 오직 그 경우에 한해 거짓임을 진술하고 있다.
ψ가 거짓 술어인 경우 “이 문장은 거짓이다.”와 같은 형식의 문장이 S를 만족하는 문장이 될 것이다. 이와 같은 노골적인 형식을 지니지는 않아도 일견 아무 문제가 없어 보이는 문장으로부터도 그와 같은 자기지시적인 문장이 유도될 수 있다. 거짓말쟁이의 역설이라는 용어를 낳게 한 크레타(Creta)인의 철학자 에피메니데스(Epimenides)가 했다고 하는 “크레타인은 모두 항상 거짓말을 한다.”는 문장이 그 한 예에 해당한다. 지금 Cret($x$)를 “$x$는 크레타 인이 한 문장이다”라는 의미의 술어라고 할 때, 에피메니데스의 진술은 다음과 같은 형태로 기호화 할 수 있다.

***e***. $(\forall x)(\mathrm{Cret}(x) \rightarrow F(x))$

에피메니데스가 말한 문장은 다음과 같은 진술의 한 예이다.

***a***. $(\forall x)(S(x) \rightarrow \psi(x))$

$a$는 즉 “모든 $S$는 ψ이다.”라는 취지의 문장이다. 그 문장에 관해 다음과 같은 사실이 성립한다고 하자.

(3) $S(a)$
(4) $(\forall x)(S(x) \wedge x \neq a \rightarrow \psi(x))$

(3)은 문장 $a$ 자체가 $S$임을 말하고 있고 (4)는 $a$ 이외에 $S$인 문장들은 모두 속성 ψ를 지니고 있음을 말하고 있다. (3)과 (4)가 성립하면 그로부터 곧장 다음이 유도된다.

(5) $\psi(a) \leftrightarrow (\forall x)(S(x) \rightarrow \psi(x))$

(5)에서 오른쪽 변의 식은 $a$가 가리키는 공식이므로 (5)는 곧 $a$가 ψ임을 주장하는 자기지시적인 문장이라고 볼 수 있다. 또한 위의 A에서 'φ'의 자리에 $a$를 그리고 φ의 자리에 $(\forall x)(S(x) \rightarrow \psi(x))$를 대입함으로써

$$(6) \qquad T(a) \leftrightarrow (\forall x)(S(x) \rightarrow \psi(x))$$

를 얻을 수 있다. (5)와 (6)으로부터 궁극적으로 다음이 얻어진다.

$$(7) \qquad T(a) \leftrightarrow \psi(a)$$

(7)에서 일종의 대각화(diagonalization)로서 술어 ψ 대신에 $F$를 대입하면 위의 (2)와 맞먹는 다음과 같은 공식을 얻을 수 있는데 이로부터 역설은 즉각적으로 초래된다. 그러한 역설이 이른바 거짓말쟁이의 역설이다.

$$(8) \qquad T(a) \leftrightarrow F(a)$$

에피메니데스의 경우에 (3)과 (4)는 각각 다음과 같은 진술이 될 것이다.

$$(3)' \qquad \mathrm{Cret}(e)$$
$$(4)' \qquad (\forall x)(\mathrm{Cret}(x) \wedge x \neq e \rightarrow F(x))$$

(3)'은 에피메니데스가 한 $e$도 크레타 인이 진술한 문장이라는 것을 의미하며 (4)'는 $e$ 이외에 크레타 인이 진술한 모든 문장은 거짓이라는 것을 말하고 있다. 따라서 $e$에 대해 다음과 같은 역설적인 결과가 나오게 된다.

$$(8)' \qquad T(e) \leftrightarrow F(e)$$

(8)'에 의하면 모든 크레타 인이 진술한 문장은 거짓이라는 문장은 실제로 그 문장을 크레타 인이 진술하였고 크레타 인이 진술한 그 밖의 다른 모든 문장이 거짓이라면 그 문장이 참일 경우 또 오직 그 경우에 한해 거짓이 된다. 이렇게 해서 $e$ 자신은 자기지시적인 문장이 아니지만 어떤 경험적인 사실이 성립할 경우 그로부터 자기지시적인 문장이 도출된다.

만일 문장 $a$나 $e$가 속한 언어가 이치 논리를 받아들인다면 (8)이나

(8)‘는 곧장 모순을 낳게 된다. 왜냐하면 (8)이나 (8)’에 의하면 $a$나 $e$는 참인 동시에 거짓이거나 참이 아님과 동시에 거짓도 아니어야 하는데 이치 논리가 성립하는 한 그 어떤 것도 불가능하기 때문이다. 그러나 그 문장들이 속한 언어에서 이치 논리가 성립하지 않으면 그로부터 반드시 모순이 초래되지 않는다. $a$나 $e$가 참도 아니지만 동시에 거짓도 아니라고 말함으로써 (8)이나 (8)‘을 수용할 수 있기 때문이다. 그러나 우리의 언어가 이치 논리가 아닌 다치 논리가 성립한다고 하더라도 (7)에서의 대각화의 과정에서 $\phi$ 대신에 $F$를 대입하는 대신에 $\neg T$를 대입함으로써 새로운 모순을 이끌어내는 것이 가능하다. (7)에서 $\phi$ 대신에 $F$를 대입하는 대신에 $\neg T$를 대입하면 다음을 얻는다.

(9) $T(a) \leftrightarrow \neg T(a)$

(9)에 의하면 $a$는 참인 동시에 참이 아니어야 하지만 이것은 다치 논리를 허용하는 상황에서도 불가능하다.

다음과 같은 거짓말쟁이 문장(liar sentence)들은 노골적으로 자기 지시적이다.

***l***. $F(l)$ ($l$은 거짓이다.)

***n***. $\neg T(n)$ ($n$은 참이 아니다.)

따라서 곧장 역설을 야기하는 문제 있는 결론 $T(l) \leftrightarrow F(l)$ 혹은 $T(l) \leftrightarrow \neg T(l)$을 얻을 수 있다. 그러나 $e$는 그 자체 자기지시적인 문장이 아니며 단지 어떤 경험적 사실, 즉 (3)‘ 특히 {4}’가 성립할 경우, 자기 지시적 문장이 귀결될 따름이다. 그러나 만일 다음과 같은 사실이 성립할 경우에는 거짓인 문장이 된다.

(10) $(\exists x)(\mathrm{Cret}(x) \wedge x \neq e \wedge T(x))$

(10)은 $e$ 이외에 크레타 인이 진술한 문장 가운데 참인 문장이 적어도 하나 존재한다는 것을 말하고 있다. 만일 (10)이 성립하면 모든 크레타 인이 진술한 모든 문장은 단순히 거짓이 되며 아무 역설도 야기되

지 않는다.

$l$과 같은 거짓말쟁이 문장으로부터 귀결되는 $T(l) \leftrightarrow F(l)$로부터 우리가 내릴 수 있는 결론은 $T(l)$과 $F(l)$이 동일한 진리값을 갖는다는 것이다. 즉 둘은 동시에 참이거나 동시에 거짓이다. 그러나 $T(l)$과 $F(l)$이 동시에 참이라는 것은 받아들일 수 없어도 이치 논리를 거부하는 한 동시에 거짓이라는 것은 받아들일 수 있다. 그것은 $l$이 참도 아니고 거짓도 아닌 경우이다. 그러나 이러한 진리치 간격(truth gap)을 허용한다고 해도 $n$과 같은 강화된 거짓말쟁이의 문장(strengthened liar sentence)이 야기하는 역설을 피해갈 수는 없을 것 같다. 왜냐하면 $n$이 거짓인 경우는 물론이지만 참도 거짓도 아닌 경우에도 $n$이 참이 아니라고 해야 한다. 그러므로 $\neg T(n)$인데 이것이 $n$이므로 $n$은 참이다. 즉 $n$이 거짓이거나 혹은 참도 거짓도 아닌 경우에도 $n$은 참이다. $n$이 참인 경우에도 $n$은 참이므로 어떤 경우에도 $n$은 참이다. 그러나 $n$이 참이라고 하면 $n$은 $\neg T(n)$이므로 $n$은 참이 아니다. 결국 $n$은 참이기도 하고 참이 아니기도 한데 이것은 모순이다.

A는 상식적인 진리 개념을 잘 반영하고 있다고 말할 수 있다. 그러나 A가 자기지시적인 문장 $l$ 혹은 $n$과 관련하여 역설을 야기한다는 것은 A에 반영된 상식적인 개념이 상당한 문제를 안고 있다는 것을 의미한다. 물론 여기서 A가 아니라 자기 지시적인 문장에 문제가 있다고 생각할 수도 있다. 그러나 $e$와 같은 비자기지시적인 문장도 어떤 경험적인 사실이 전제될 경우 역설을 야기한다는 점에 비추어 자기 지시성에서 역설의 원인을 찾는 진단은 받아들이기가 어렵다. 이러한 이유에서 타르스키는 A에서 반영되는 개념은 진리의 개념으로서 충분하지 않다고 생각했다. 그는 역설을 야기하지 않고는 A를 만족시키는 진리 개념을 정의하기는 불가능하다고 보았다. 그는 통상적인 논리 법칙을 버리지 않는 한, 어떤 언어가 그 언어에 속하는 모든 표현을 지시하는 이름을 포함하고 또한 그 언어에 속하는 문장들을 지시하는 것으로서 '참'과 같은 의미론적 용어를 포함한다면, 다시 말해 그가 말하는 "의미론

적으로 폐쇄된”(semantically closed) 상태에 있는 한, 자기지시적인 거짓말쟁이의 문장과 같은 것을 통해 역설이 야기되는 것은 피할 수 없다고 생각했다.[1)]

타르스키는 어떤 언어에서의 참을 의미하는 술어가 그 언어에는 속하지 않더라도 그 언어를 기술하는 언어, 즉 메타언어(meta-language)에는 속해야 한다고 생각했다. 다시 말해 대상언어를 $\mathcal{L}$이라고 하면 ‘$\mathcal{L}$에서의 참’(true-in-$\mathcal{L}$)은 그것을 기술하는 메타언어에 속해야 한다는 것이다. 그러한 진리 개념을 $T_{\mathcal{L}}$이라고 할 때, 타르스키는 일상적인 언어가 아닌 형식 언어를 대상으로 그 개념의 외연을 규정함으로써 $T_{\mathcal{L}}$을 정의하려는 시도를 하고 있다. 타르스키는 형식 언어에 대한 어떤 진리 정의도 위의 A와 유사한 조건을 충족시켜야 한다고 생각했다. 대상언어와 메타언어를 확연히 구분하려는 정신에 비추어 볼 때 A는 다음과 같이 정식화 되어야 한다. 즉 $\mathcal{L}$의 모든 문장 $\phi$에 대해,

**T**. $r_{\phi} \in T_{\mathcal{L}} \Leftrightarrow \phi'$

위에서 $r_{\phi}$는 $\mathcal{L}$의 문장 $\phi$를 가리키는 메타언어의 이름이고 $\phi$‘는 메타언어에서의 $\phi$의 번역에 해당하며, $\Leftrightarrow$는 쌍조건을 의미하는 메타언어의 기호이다.[2)] 타르스키는 위와 같은 그가 말하는 규약(convention) **T**를 충족한다는 의미에서 대상 언어에 대한 진리 정의는 실질적으로 적합해야(materially adequate) 한다고 주장했다. 규약 **T**에 비추어 볼 때 메타언어는 대상언어의 모든 문장에 대한 번역을 포함하고 있어야 한다. 그러나 대상언어는 메타언어에 속하는 $T_{\mathcal{L}}$에 해당하는 표현을 지닐 수 없다. 따라서 메타언어가 대상언어의 진리 개념을 만족할만하게 정의할 수 있기 위해서는 그 표현력이 대상언어보다 풍부하지 않으면 안 된다.

타르스키는 형식 언어의 진리 개념에 대한 정의를 이른바 모델

---

1) Tarski[1944], 20 쪽.

2) Tarski[1935], 187-8 쪽. 타르스키는 $r\phi$이 대상언어에 속하는 문장을 가리키는 ‘구조 기술적’{structural- descriptive) 이름이어야 한다고 말하고 있지만 그가 말하는 구조 기술적 이름이 무엇인지는 여기서 중요하지 않다.

(model)을 토대로 내리고 있다.[3] 타르스키의 모델 $\mathcal{M}$은 그것의 영역이라고 불리는 공이 아닌 집합 $M$과 $\mathcal{L}$의 각 비논리 상항에 적절한 대상을 대응시켜 주는 해석 함수 $v$로 이루어진다. 따라서 그것을 순서쌍 $<M, v>$로 나타낼 수 있다. $\mathcal{L}$의 비논리 상항은 물론 개별 상항(individual constant), 술어(predicate), 함수기호(function symbols)들로 이루어지며 기본(primitive) 논리 상항은 각기 부정(negation), 선언(disjunction), 존재양화(existential quantification)를 의미하는 ¬, ∨, ∃로 생각할 수 있다.

해석 함수 $v$는 $\mathcal{L}$의 각 $n$항의 술어에 집합 $M$에서의 $n$항 관계를 대응시키고 각 개별 상항에 $M$의 원소를 대응시킨다. 즉 개별 상항 $a$에 대해 $v(a) \in M$이고 $n$항의 술어 $P^n$에 대해 $v(P^n) \subset M^n$. 또한 $n$항의 함수기호에 대해서는 $M$에서의 $n$항의 함수를 대응시킨다. 이러한 대응 관계에 의해 $\mathcal{L}$의 각 항(term) $t$에 대해 $M$의 한 원소를 할당할 수 있다. 그 값을 $v(t)$라고 하자. 타르스키는 이러한 대응관계를 바탕으로 각 문장에 대해 '$\mathcal{M}$에서의 참'을 다음과 같이 귀납적으로 정의한다. '$\mathcal{M}$에서의 참'을 기호 '$\mathcal{M} \vDash$'로, 그리고 '$\mathcal{M}$에서의 참이 아님을' 기호 ''$\mathcal{M} \nvDash$'로 '나타내면,

1. $n$항의 술어 $P^n$과 항 $t_i, \cdots, t_n$에 대해, $\mathcal{M} \vDash P^n t_i \cdots t_n \Leftrightarrow <v(t_1), \cdots, v(t_n)> \in v(P^n)$.

2. 각 문장 $\phi$에 대해 $\mathcal{M} \vDash \neg\phi \Leftrightarrow \mathcal{M} \nvDash \phi$.

3. 각 문장 $\phi$와 $\psi$에 대해 $\mathcal{M} \vDash \phi \vee \psi \Leftrightarrow \mathcal{M} \vDash \phi$이거나 $\mathcal{M} \vDash \psi$.

4. 각 문장 $\phi$와 변항 $x$에 대해 $\mathcal{M} \vDash (\exists x)\phi(x) \Leftrightarrow$ 어떤 모델 $\mathcal{M}_a$에 대해 $\mathcal{M}_a \vDash \phi(a)$

4에서 $a$는 $\phi(x)$에는 등장하지 않는 개별 상항이며 모델 $\mathcal{M}_a$는 영역이

3) 이러한 이유로 논리에 관한 근래의 저술에서는 '언어 $\mathcal{L}$에서의 참'(true-in-$\mathcal{L}$)이라는 용어보다는 '모델 $\mathcal{M}$에서의 참'(true-in-model M)이라는 용어를 선호하는 경향이 있다.

$\mathscr{M}$과 동일하고 그것을 이루는 또 다른 원소인 해석 함수 $v_a$는 개별 상항 $a$만을 제외하고는 $\mathscr{L}$의 모든 비논리 상항에 대해 $v$와 동일한 값을 할당한다. 타르스키는 위의 정의에 의해 집합 $\{\phi \mid \mathscr{M} \vDash \phi\}$가 일의적으로 결정되며 그 집합을 진리 개념 $T_{\mathscr{L}}$로 정의할 경우, 즉 모든 $\phi$에 대해 $\mathscr{M} \vDash \phi \Leftrightarrow r_{\phi} \in T_{\mathscr{L}}$이라고 할 경우 위의 규약 T가 충족된다는 것을 보이고 있다.

$\mathscr{L}$이 산수학(arithmetic)을 전개할 정도로 풍부한 언어일 경우 괴델(K. Gödel)이 한 방식에 따라 $\mathscr{L}$의 각 문장에 대해 하나의 숫자(numeral)을 대응시킬 수 있으며 그것을 $\mathscr{L}$에서의 그 문장의 이름으로 간주할 수 있다. 지금 $\phi$의 $\mathscr{L}$에서의 이름을 '$\phi$'라고 하자. 우리는 $\mathscr{M}$의 영역 $M$이 $\mathscr{L}$의 모든 문장 $\phi$에 대해 $v('\phi')$를 포함하는 것으로 가정할 수 있으며 $v('\phi')$를 $r_{\phi}$로 잡을 수도 있다. 만일 다음을 성립하는 $\mathscr{L}$의 일항 술어 $Tr$이 존재한다면 그것을 우리는 진리 개념 $T_{\mathscr{L}}$를 M에서 정의하는 술어라고 부를 수 있을 것이다. 모든 $\phi$에 대해

$$\text{(11)} \qquad \mathscr{M} \vDash Tr('\phi) \Leftrightarrow r_{\phi} \in T_{\mathscr{L}} \;\; (v('\phi') \in T_{\mathscr{L}})$$ [4)]

모든 $\phi$에 대해 $\mathscr{M} \vDash \phi \Leftrightarrow r_{\phi} \in T_{\mathscr{L}}$인 것에 비추어 (11)로부터 곧장 다음을 얻을 수 있다.

$$\text{(12)} \qquad \mathscr{M} \vDash Tr('\phi) \Leftrightarrow \mathscr{M} \vDash \phi$$

'$\mathscr{M} \vDash$'의 정의에 비추어 (12)는 다음이 성립할 경우 또 오직 그 경우에 성립할 것이다. 즉 $\mathscr{L}$의 모든 문장 $\phi$에 대해,

$$\mathbf{T}_{\mathscr{L}}. \qquad \mathscr{M} \vDash (Tr('\phi) \leftrightarrow \phi)$$

위의 관계를 우리는 대상언어 $\mathscr{L}$에서 메타언어에서의 규약 **T**를 표현하는 관계라고 생각할 수 있다. 즉 메타언어의 술어 $T_{\mathscr{L}}$을 정의하는 술

4) 일반적으로 메타언어에 속하는 $n$-항의 관계 R에 대해 $\langle v(a_1), \cdots, v(a_n)\rangle \in \mathrm{R} \Leftrightarrow \mathscr{M} \vDash \psi(a_1, \cdots, a_n)$인 술어 $\psi$가 대상언어 $\mathscr{L}$에 존재할 경우 R은 $\mathscr{M}$에서 정의가능하다고 말하고 $\psi$를 R을 정의하는 술어라고 한다.

어가 대상언어 $\mathscr{L}$에서 존재할 경우 규약 **T**를 $\mathscr{L}$에서 표현하는 일도 가능해진다. 여기에 $\mathscr{L}$의 임의의 일항 술어 $\psi$에 대해 자신이 $\psi$임을 말하는 자기 지시적인(self-referring) 문장이 존재한다고 하자. 그러한 자기 지시적 문장 $S_\psi$는 다음을 만족할 것이다.

$$(13) \qquad \mathscr{M} \vDash (\psi('S_\psi') \leftrightarrow S_\psi)$$

(13)을 만족하는 $S_\psi$를 흔히 술어 $\psi$의 부동점(fixed point)이라고 부른다. 위의 지금 진리 개념 $T_{\mathscr{L}}$을 $\mathscr{M}$에서 정의하는 술어 $Tr$이 $\mathscr{L}$에 존재한다면 $\neg Tr$도 존재할 것이다. 그러한 술어의 부동점을 $\lambda$라고 하면 $\lambda$에 대해 다음이 성립할 것이다.

$$(14) \qquad \mathscr{M} \vDash (\neg Tr('\lambda') \leftrightarrow \lambda)$$

또한 관계 $T_{\mathscr{L}}$로부터

$$(15) \qquad \mathscr{M} \vDash (Tr('\lambda') \leftrightarrow \lambda)$$

(14)와 (15)로부터 곧장 형식화한 언어 $\mathscr{L}$에서의 거짓말쟁이의 역설에 해당하는 결론 $\mathscr{M} \vDash (\neg Tr('\lambda') \leftrightarrow Tr('\lambda'))$를 얻을 수 있는데 이것은 $\mathscr{L}$이 모순된 체계임을 말해준다. 그런데 $\mathscr{L}$은 형식 언어체계로서 모순이 없게 만들 수 있다. 그렇다면 진리 개념 $T_{\mathscr{L}}$을 정의하는 술어 $Tr$이 $\mathscr{L}$에 존재하지 않거나 혹은 모든 일항 술어에 대해 그 부동점이 존재하지 않는다고 생각해야 한다. 그런데 괴델이 그의 정리를 증명하는 과정에서 보인 것처럼 $\mathscr{L}$이 산수학을 전개할 정도로 풍부한 언어라면 그것의 모든 일항 술어에 대해 부동점이 존재한다. 이러한 부동점 정리에 비추어 볼 때, $\mathscr{L}$에서는 진리 술어 $T_{\mathscr{L}}$을 정의하는 술어가 존재하지 않는다는 결론을 내려야 한다.

타르스키는 그가 다루는 형식 언어가 A에 해당하는 성질을 갖는 술어를 그것에서의 진리 개념을 정의하려는 언어에서 정의할 수 없도록 하는 전략을 채택함으로써 거짓말쟁이의 역설을 피해 가고 있다. A와 유사한 규약 **T**를 만족하는 술어 $T_{\mathscr{L}}$이 존재하지만 그 술어는 $\mathscr{L}$이 아닌

그것의 메타언어에 속하는 술어이다. 이것은 곧 문제의 메타언어에서의 진리 개념을 정의하자면 또 그것의 메타언어로 물러나야 한다는 것을 의미한다. 이렇게 해서 메타언어에서의 진리 술어 $T_{M}$을 정의할 수 있지만 그 술어는 사실 이제는 새로운 대상언어가 된 메타언어 M에는 속하지 않는다. 이렇게 해서 타르스키의 경우 무한히 거슬러 올라가는 진리 술어의 계열 $T_{\mathcal{L}}$, $T_{M}$, …을 얻게 된다.

## 2. 부분 모델(partial model)과 진리치 간격

타르스키는 거짓말쟁이의 역설을 피하는 대가로 대상언어 자체에서 그 언어에서의 진리를 표현하는 가능성을 봉쇄하는 길을 취하고 있다. 대상언어에서의 진리 개념은 오직 상위의 메타언어에서만 표현할 수 있는데 그러기 위해서는 또한 대상언어의 모든 문장을 메타언어로 번역하는 것이 가능해야 한다. 콰인(Willard, V, O. Quine)이 제기한 번역의 불확정성 논제, 즉 대상언어로부터 메타언어로의 단일한 올바른 번역이 가능한가 하는 문제에 비추어 볼 때, 이러한 타르스키의 선택이 일정한 제약을 갖는다는 것은 분명해 보인다. 그러나 어떤 언어에서의 진리 개념을 그 언어 안에서 표현할 수 없다는 선택이 가져다주는 제약은 더욱 분명하다. 이러한 제약으로 인해 한 언어의 진리 개념을 정의하기 위해서는 반드시 보다 표현력이 풍부한 상위의 언어를 요구하게 되었으며 그로 인해 진리 개념은 계층을 달리하는 각 언어에서의 진리 개념으로 분화될 수밖에 없다. 이러한 상황에서 전절에서의 (12) 혹은 관계 $\mathbf{T}_{\mathcal{L}}$과 같은 의미에서 대상 언어 $\mathcal{L}$에서의 진리 개념을 표현하는 진리 술어를 $\mathcal{L}$에서 정의할 가능성은 없는가 하는 질문을 제기할 수 있다. 크립키(S. Kripke)는 그의 Kripke[1975]에서 이러한 물음에 대해 제한적으로 긍정적인 답변을 제시하고 있다.

크립키의 목표는 언어 $\mathcal{L}$에서 위의 (12)와 같은 관계를 만족하는 술

어를 정의하는 것이다. 그러한 술어를 정의함에 있어 크립키는 고전적인 모델 대신에 부분 모델을 사용하고 있다. 크립키의 부분 모델은 타르스키가 사용한 고전적인 모델과는 달리 모델의 원소인 해석 함수 $v$가 술어에 대해 서로 겹치지 않는 두 개의 집합을 대응시킨다. 즉, $E$, $A\subset M^n$이고 $E\cap A=\varnothing$인 어떤 $E$와 $A$에 대해 $v(P^n)=<E, A>$. 크립키는 집합 $E$를 술어 $P^n$의 외연(extension), A를 반외연(anti-extension)이라고 부르고 있다. 외연은 그것에 대해 술어 $P^n$이 성립하는 원소들로 이루어진 집합, 그리고 반외연은 그것에 대해 술어 $P^n$이 성립하지 않는 원소로 생각할 수 있다. 이러한 부분 모델을 바탕으로 크립키는 그 모델에서 참을 의미하는 '$\mathscr{M}\vDash$'과 거짓임을 의미하는 '$\mathscr{M}\Vdash$'을 정의하고 있다. 그러한 정의는 다음과 같이 진행된다.

1. 어떤 문장 $\phi$에 대해서도 $\mathscr{M}\vDash\phi$와 $\mathscr{M}\Vdash\phi$가 동시에 성립하지 않는다.
2. $v(P^n)=<E, A>$.인 $n$항의 술어 $P^n$과 항 $t_i$, …, $t_n$에 대해,
   $<v(t_1), \cdots, v(t_n)>\in E$인 경우 $\mathscr{M}\vDash P^n t_i\cdots t_n$이고
   $<v(t_1), \cdots, v(t_n)>\in A$인 경우 $\mathscr{M}\Vdash P^n t_i\cdots t_n$.
3. 각 문장 $\phi$에 대해,
   $\mathscr{M}\vDash\neg\phi \Leftrightarrow \mathscr{M}\Vdash\phi$,
   $\mathscr{M}\Vdash\neg\phi \Leftrightarrow \mathscr{M}\vDash\phi$
4. 각 문장 $\phi$와 $\psi$에 대해,
   $\mathscr{M}\vDash\phi\vee\psi \Leftrightarrow \mathscr{M}\vDash\phi$이거나 $\mathscr{M}\vDash\psi$,
   $\mathscr{M}\Vdash\phi\vee\psi \Leftrightarrow \mathscr{M}\Vdash\phi$이고 $\mathscr{M}\Vdash\psi$,
5. 각 문장 $\phi$와 변항 $x$에 대해,
   $\mathscr{M}\vDash(\exists x)\phi(x) \Leftrightarrow$ 어떤 모델 $\mathscr{M}_a$에 대해 $\mathscr{M}_a\vDash\phi(a)$
   $\mathscr{M}\Vdash(\exists x)\phi(x) \Leftrightarrow$ 모든 모델 $\mathscr{M}_a$에 대해 $\mathscr{M}_a\Vdash\phi(a)$

5의 경우에도 $\mathscr{M}_a$는 영역이 $\mathscr{M}$과 동일하고 그것을 이루는 함수 $v_a$는 $\phi(x)$에는 등장하지 않는 개별 상항 $a$만을 제외하고는 $\mathscr{L}$의 모든 비논리

상항에 대해 $v$와 동일한 값을 할당하는 부분 모델이다.

위의 1-5에 의해 $\mathcal{L}$의 각 문장 ϕ에 대해서는 (ⅰ) $\mathcal{M}\vDash\phi$이거나 (ⅱ) $\mathcal{M}\,\Vdash\phi$이거나 혹은 (ⅲ) $\mathcal{M}\nvDash\phi$이고 또한 $\mathcal{M}\nVdash\phi$이거나 세 가지 가운데 하나가 성립한다. 첫 번째 경우는 ϕ가 $\mathcal{M}$에서 참인 경우이고 두 번째는 거짓인 경우, 세 번째는 참도 거짓도 아닌 경우이다. 이렇게 해서 부분 모델은 하나의 문장이 참도 거짓도 아닌 경우, 즉 진리치가 정의되지 않는 경우를 인정하게 된다.[5)]

동일한 언어 $\mathcal{L}$의 두 모델 $\mathcal{M}=<M, v>$와 $\mathcal{M}'=<M', v'>$에 대해 $M=M'$이고 술어를 제외한 $\mathcal{L}$의 모든 비논리상항 $c$에 대해 $v(c)=v'(c')$이며 $\mathcal{L}$의 각 술어 $P$에 대해 $v(P)\leq v'(P)$라고 하자. 즉, $v(P)=<E, A>$, $v'(P)=<E', A'>$라고 할 때, $E\subset E'$, $A\subset A'$라고 하자. 이 경우 $\mathcal{M}'$을 $\mathcal{M}$의 연장 모델(extension of M)이라고 한다. $\mathcal{M}'$을 $\mathcal{M}$의 연장 모델이라는 것을 기호로 $\mathcal{M}\leq\mathcal{M}'$으로 나타내기로 하자. 문장의 복잡성을 이용한 통상적인 귀납에 의해 위의 1-5에 의해 정의되는 모델에서의 참과 거짓에 대해 다음이 성립함을 쉽게 증명할 수 있다.

- $\mathcal{M}\leq\mathcal{M}'$이면 모든 ϕ에 대해 $\mathcal{M}\vDash\phi$이면 $\mathcal{M}'\vDash\phi$이고 $\mathcal{M}\Vdash\phi$이면 $\mathcal{M}'\Vdash\phi$.

즉, 위의 정의 1-5는 위와 같은 의미에서 모델의 단조성(monotonicity)을 보장한다.

앞서 언급한 대로 크립키의 목표는 대상 언어 $\mathcal{L}$에 속하는 것으로서 (12)와 같은 조건을 만족하는 진리 술어를 정의하는 것이다. 그런데 우리의 모델 $\mathcal{M}$은 부분 모델로서 어떤 문장에는 참을, 그리고 다른 문장에는 거짓을 할당한다는 사실에 비추어 우리가 정의하려는 진리 술어 τ는 다

5) 특히 선언문과 존재 양화 문장에 진위를 배당하는 조건 4와 5를 오른쪽 변에 정의되지 않는 문장이 하나라고 존재할 경우 선언문과 존재 양화 문장은 진위가 정의되지 않는다는 것으로 약화시킬 수도 있다. 그런 식으로 4-5를 약화시킨 진위 배당 방식을 약한 클리니(Kleene) 진위 배당, 그리고 1-5와 같은 진위 배당을 강한 클리니 배당이라고 한다. 여기서는 강한 클리니 진위 배당만을 고찰할 것이지만 이하의 논의는 약한 경우에도 비슷한 방식으로 성립한다.

음과 같은 조건을 만족해야 할 것이다. 즉, 모든 $\mathcal{L}$의 문장 $\phi$에 대해,

$$(16)\qquad \mathcal{M} \models \tau(`\phi') \Leftrightarrow \mathcal{M} \models \phi,$$
$$\mathcal{M} \mathrel{|}\!\!\models \tau(`\phi') \Leftrightarrow \mathcal{M} \mathrel{|}\!\!\models \phi$$

이제 $\mathcal{L}$에 새로운 일항 술어 $T$를 도입하여 그 언어를 확장하기로 하고 그렇게 확장된 언어를 $\mathcal{L}+T$로 표기하기로 하자. 지금 편의상 $\mathcal{L}+T$의 부분 모델 $\mathcal{M}'$의 영역이 언어 $\mathcal{L}+T$의 모든 문장을 원소로 포함한다고 가정해도 무방하다. 확장된 언어 $\mathcal{L}+T$의 부분 모델 $\mathcal{M}'=<M', v'>$은 물론 기존의 $\mathcal{L}$의 모델 $\mathcal{M}=<M, v>$와는 다를 것이다. 그러나 다음과 같은 조건을 만족시킬 수 있다. $M=M'$이고 $\mathcal{L}$의 모든 비논리상항 $c$에 대해 $v(c)=v'(c')$. 즉 $\mathcal{M}$과 $\mathcal{M}'$은 영역이 동일하며 술어 $T$를 제외한 $\mathcal{L}+T$의 모든 비논리상항을 동일한 방식으로 해석한다. 이러한 모델 $\mathcal{M}'$을 $\mathcal{M}$의 확장 모델(expansion of $\mathcal{M}$)이라고 한다.

$\mathcal{M}$의 확장 모델 $\mathcal{M}'=<M', v'>$은 $\mathcal{M}$과는 달리 $T$에도 외연과 반외연을 할당함으로써 $T$를 해석한다. 만일 $v'(T)=<\{\phi \mid \mathcal{M} \models \phi\}, \{\phi \mid \mathcal{M} \mathrel{|}\!\!\models \phi\} \cup N>$와 같이 해석한다면 그것은 $T$를 '$\mathcal{L}$에서의 참' 혹은 '모델 $\mathcal{M}$에서 참'의 개념으로 해석한다는 말이 될 것이다. (여기서 $N$은 $\mathcal{L}+T$의 문장이 아닌 논리식들의 집합이다.) 왜냐하면 $\phi$가 $\mathcal{M}$에서 참인 경우 또 오직 그 경우에 한해 $T$의 외연에 속할 것이며 $\phi$가 $\mathcal{M}$에서 거짓이거나 혹은 문장이 아닐 경우 $T$의 반외연에 속할 것이기 때문이다. 그러나 그러한 $\mathcal{M}'$에 대해 (16)은 성립하지 않는다. 크립키의 생각은 $\mathcal{M}$의 확장 모델 $\mathcal{M}'$을 단계적으로 연장해 감으로써 궁극적으로 (16)을 만족하는 모델을 얻는다는 것이다. 만일 그러한 모델을 얻는다면 그 모델에 의해 정의된 술어 $T$가 바로 대상 언어 $\mathcal{L}+T$에 속하면서도 그 언어에서의 진리를 의미하는 술어가 될 것이다.

크립키가 생각하는 $\mathcal{M}$의 확장 모델들의 계열 $\mathcal{M}_0=<M, v_0>$, $\mathcal{M}_1=<M, v_1>$, …은 모두 해석 함수 $v$가 $T$에 할당하는 값에 대해서만 차이가 있으므로 그 모델들을 정의하기 위해서는 해석 함수 $v_0(T)$, $v_1(T)$, …등을

정의하는 것으로 충분하다. 크립키는 $v_0(T)$, $v_1(T)$, …를 다음과 같이 귀납적인 방식으로 정의할 것을 제안한다.[6)]

$$v_0(T) = <\varnothing,\ \varnothing>$$
$$v_1(T) = <E_1,\ A_1> = <\{\phi \mid \mathscr{M}_0 \vDash \phi\},\ \{\phi \mid \mathscr{M}_0 \nvDash \phi\} \cup N>$$
$$\vdots$$
$$v_{j\alpha+1}(T) = <E_{\alpha+1},\ A_{\alpha+1}> = <\{\phi \mid \mathscr{M}_{\alpha j} \vDash \phi\},\ \{\phi \mid \mathscr{M}_{\alpha} \nvDash \phi\} \cup N>$$
$$\vdots$$

위의 정의에서 보는 것처럼 α+1-단계에서의 확장 모델 $\mathscr{M}_{\alpha+1}$의 해석함수에 의한 술어 $T$의 외연은 α-단계에서의 모델에 의해 참과 거짓으로 분류되는 문장들의 집합으로 결정된다. 그러므로 $\mathcal{L}+T$의 모든 문장 $\phi$에 대해, $\phi \in E_{\sigma+1} \Leftrightarrow \mathscr{M}_{\alpha} \vDash \phi$, $\phi \in A_{\sigma+1} \Leftrightarrow \mathscr{M}_{\alpha} \nvDash \phi$이거나 $\phi$는 비문장(non-sentence). 따라서 모델의 정의에 의해 $\mathscr{M}_{\alpha+1} \vDash T('\phi') \Leftrightarrow \mathscr{M}_{\alpha} \vDash \phi$, $\mathscr{M}_{\alpha+1} \nvDash T('\phi') \Leftrightarrow \mathscr{M}_{\alpha} \nvDash \phi$이거나 $\phi$는 비문장.

명백히 $v_0(T) \leq v_1(T)$. 따라서 $\mathscr{M}_0 \leq \mathscr{M}_1$. 이제 $v_n(T) \leq v_{n+1}(T)$이라고 하자. 그러면 $\mathscr{M}_n \leq \mathscr{M}_{n+1}$. 따라서 단조성에 의해 $\mathcal{L}+T$의 모든 문장 $\phi$에 대해 $\mathscr{M}_n \vDash \phi$이면 $\mathscr{M}_{n+1} \vDash \phi$이고 $\mathscr{M}_n \nvDash \phi$이면 $\mathscr{M}_{n+1} \nvDash \phi$. 이로부터 $\phi \in E_{n+1}$이면 $\phi \in E_{n+2}$이고 $\phi \in A_{n+1}$이면 $\phi \in A_{n+2}$라는 결과를 얻을 수 있다. 따라서 $v_{n+1}(T) \leq v_{n+2}(T)$. 이제 $v_0(T) \leq v_1(T)$이고 모든 $n$에 대해 $v_n(T) \leq v_{n+1}(T)$이면 $v_{n+1}(T) \leq v_{n+2}(T)$이므로 수학적 귀납법에 의해 모든 수 $n$에 대해 $v_n(T) \leq v_{n+1}(T)$.[7)] 따라서 $\mathscr{M}_n \leq \mathscr{M}_{n+1}$.

위와 같이 정의된 $v_0$, $v_1$, …는 단조적으로 증가하는 열을 이루지만 특기할만한 점은 $\lambda < \xi$인 모든 $\xi$에 대해 $v_\lambda = v_\xi$인 서수 $\lambda$가 존재한다는 것이다.[8)] 그러한 $v_\lambda$에 의해 주어지는 $v_\lambda(T) = <E_\lambda,\ A_\lambda>$를 부동점(fixed

---

6) 크립키는 자연수만이 아니라 초한서수 α에 대해서도 $v_\alpha(T)$를 정의하는 방법을 제시하고 있다. Kripke[1975], 704 쪽.

7) 수학적 귀납법에 의하면 모든 자연수 혹은 유한수에 대해서만 이 관계가 성립한다. 그러나 크립키는 초한 서수들에 대해서도 이러한 관계가 성립한다고 말하고 있다. Kripke[1975], 704 쪽 참조.

point)이라고 부른다. $<E_\lambda, A_\lambda>$를 부동점이라고 하면 특히 $<E_\lambda, A_\lambda>=<E_{\lambda+1}, A_{\lambda+1}>$. 즉, $E_\lambda=E_{\lambda+1}$, $A_\lambda=A_{\lambda+1}$. 부분 모델에 의함 참 거짓의 정의에 의해 모든 문장 $\phi$에 대해, $\phi\in E_\lambda \Leftrightarrow \mathscr{M}_\lambda\vDash T('\phi')$. 또한 $E_\lambda=E_{\lambda+1}$이므로 $\phi\in E_\lambda \Leftrightarrow \phi\in E_{\lambda+1}$. 또한 $E_{\lambda+1}$의 정의에 의해, $\phi\in E_{\lambda+1} \Leftrightarrow \mathscr{M}_\lambda\vDash\phi$. 따라서 궁극적으로 $\mathscr{M}_\lambda\vDash\phi \Leftrightarrow \mathscr{M}_\lambda\vDash T('\phi')$. 비슷한 방식으로 모든 문장 $\phi$에 대해, $\mathscr{M}_\lambda\,\#\!\vDash\phi \Leftrightarrow \mathscr{M}_\lambda\,\#\!\vDash T('\phi')$. 이러한 결과는 부동점에서 해석되는 $T$에 대해서는 위의 조건 (16)이 성립한다는 것을 보여주는 것이다. 따라서 그렇게 해석된 술어 $T$는 대상언어 $\mathscr{L}+T$에 *속하면서도* 그 언어에 속하는 모든 문장의 진리 개념을 표현하는 술어라고 할 수 있다.

부동점에 대해 특기할만한 것은 첫째로 외연 $E_\lambda$와 반외연 $A_\lambda$가 서로 겹치지는 않지만(disjoint) 그러나 그것을 합칠 때 완전히 영역과 동일하지 않기 때문에 그 둘 어디에도 속하지 않는, 따라서 진리 값이 정의되지 않는 문장을 허용한다는 것이다. 이러한 의미에서 크립키의 의미론은 진리치 간격(truth-value gap)을 허용하고 있다. 또 한 가지는 부동점이 하나가 아니라는 것이다. 위에서 우리는 $v_0(T)$를 이루는 외연과 반외연 $E_0$, $A_0$이 모두 공집합인 경우부터 시작했다. 그러나 $A_0$은 공집합이지만 $E_0$은 다음과 같은 문장 $\psi$ 하나로 이루어진 경우를 생각할 수 있다.

(17) $\psi = T('\psi')$

(17)은 $\psi$는 자신이 참임을 말하는 진리 문장(truth-teller)임을 의미하고 있다. $\psi\in E_0$이므로 $\mathscr{M}_0\vDash T('\psi')$. 따라서 (17)에 의해 $\mathscr{M}_0\vDash\psi$. 그러므로 $E_1$의 정의에 의해 $\psi\in E_1$. 그러므로 여전히 $v_0(T)\leq v_1(T)$. 따라서 $v_0(T)=<\{\psi\}, \varnothing>$으로부터 시작된 부동점도 존재해야 하며 이것은 위에서 정의한 부

8) 그러한 서수가 존재하리라는 것은 형식적으로 엄밀하게 증명하지 않더라도 다음과 같이 직관적으로 이해할 수 있다. $\mathscr{L}+T$의 문장들은 셀 수 있을 만큼 무한한 집합을 이룬다. 그러한 문장들의 일부는 모델 $\mathscr{M}_0$에 의해 참 혹은 거짓이 결정되며 일부는 결정되지 않는다. 그런데 열 $v_0$, $v_1$, …는 단조적으로 증가하기 때문에 모델 $\mathscr{M}_0$, $\mathscr{M}_1$, …에 의해 참 거짓이 결정되는 문장들은 계속적으로 증가한다. 따라서 그 열이 초한적으로 무한하다고 하면 어느 단계에서는 모든 문장들이 그 모델에 의해 참, 거짓이 결정될 것이다. Kripke[1975], 같은 곳 참조

동점보다 클 것이다. 위에서 우리가 구한 부동점은 모든 부동점보다 작다는 의미에서 최소의 부동점(the smallest fixed point)이라고 할 수 있다. 이러한 최소의 부동점에서 외연이나 반외연에 속하는 문장을 크립키는 기초가 있는(grounded) 문장이라고 하고 그렇지 않은 문장을 기초가 없는(ungrounded) 문장이라고 부르고 있다. 위의 (17)에서의 ψ와 같은 것은 기초가 없는 문장이다. ψ는 그러나 위에서 본 것처럼 최소가 아닌 어떤 부동점에서는 참의 진리 값을 갖는다.

위에서는 ψ를 처음 $T$의 외연에 포함시키는 해석 함수로부터 시작했지만 반외연에 포함시키는 해석 함수로부터 시작할 수도 있다. 이 경우 ψ에 거짓의 진리 값을 부여하는 부동점이 얻어질 것이다. 이처럼 자신이 참임을 말하는 진리 문장은 기초가 없지는 않지만 참으로도 거짓으로도 해석할 수 있다는 점에서 병적인 문장은 아니라고 말할 수 있다. 그러나 다음과 같은 문장 χ는 기초가 없으면서도 어떤 부동점에도 속하지 않는다.

(18) $\chi = \neg T('\chi')$

χ는 바로 자신이 거짓임을 말하는 거짓말쟁이의 문장이다. 이 문장은 명백히 최소의 부동점의 외연에도 반외연에도 속하지 않는다. 그것이 어떤 부동점에서 진리 값을 가지기 위해서는 $T$를 처음에 $v_0(T) = <\{\chi\}, \varnothing>$와 같이 해석하거나 혹은 $v_0(T) = <\varnothing, \{\chi\}>$와 같이 해석하는 것이 가능해야 한다. 그러나 그 어떤 해석도 가능하지 않다. 예를 들어 $\chi \in E_0$라고 하면 $\mathscr{M}_0 \models T('\chi')$. 따라서 $\mathscr{M}_0 \not\models \neg T('\chi')$. (18)에 의해 $\mathscr{M}_0 \not\models \chi$. 그러므로 $\chi \in A_1$. 부분 모델의 정의에 의해 $E_1$과 $A_1$은 겹치지 않으므로 $\chi \notin E_1$. 그러므로 $E_0 \not\subset E_1$이 아니며 따라서 $v_0(T) = <\varnothing, \{\chi\}>$로부터 시작하여 얻어지는 부동점은 존재하지 않는다. 이러한 의미에서 χ는 기초가 없으면서도 어떤 식으로건 진리 값을 부여할 수 없는 병적인 문장이라고 할 수 있다.

## 3. 의미론적 역설의 해결과 진리 정의

거짓말쟁이의 역설은 어떤 언어 $\mathscr{L}$에서 (18)에서와 같은 자기 지시적인 문장의 존재가 경우에 따라서는 경험적인 방식으로 확인 가능하다는 것에서 시작한다. 만일 (18)에서와 같은 술어 $T$가 다음과 같은 조건을 만족하면 모순이 발생한다는 것이 바로 거짓말쟁이의 역설이다. 즉 모든 $\mathscr{L}$의 문장 $\phi$에 대해,

$$(19) \qquad \mathscr{M} \vDash (T(‘\phi’) \leftrightarrow \phi)$$

(19)로부터 얻어지는 $\mathscr{M} \vDash (T(‘\chi’) \leftrightarrow \chi)$는 (18)로부터 얻어지는 $\mathscr{M} \vDash (\chi \leftrightarrow \neg T(‘\chi’))$와 결합되면 모순을 발생시킨다. 그러한 모순은 (18)과 (19)로부터 얻어지므로 그 가운데 적어도 하나에서 모순을 발생시킨 원인을 찾아내야 한다.

(18)은 $\chi$가 자신이 $T$가 아님을 주장하는 자기지시적인 문장이다. 자신에게 어떤 술어가 적용된다고 말하거나 적용하지 않는다고 말하는 자기지시적인 문장은 흔히 있다. 예를 들어 "이 문장은 영어로 쓰여 있지 않다."는 문장은 아무 문제될 것이 없으며 실제로 참이라고 생각될 수 있다. (18)과 같이 정의되는 $\chi$에 특별한 문제점을 인정할 수 없다면 역설의 해결을 위해 눈을 돌려야 할 것은 (19)이다. (18)은 문장 $\chi$를 정의한다고 생각되는 반면에 (19)는 $T$의 의미를 규정해 주는 것으로 생각된다. 따라서 (19)에 손을 댐으로써 역설을 해결한다는 것은 폴락(J. Pollack)[9]이 말한 것처럼 $T$의 의미를 다시 분석한 다음 그러한 분석을 토대로 한 새로 규정된 $T$의 의미에 비추어 $\chi$와 같은 거짓말쟁이 문장이 더 이상 역설을 야기하지 않는다는 것을 보이는 것이다. 위에서 설명한 타르스키나 크립키의 진리 정의도 이러한 의미에서의 역설의 해결을 지향하고 있다고 말할 수 있다. 타르스키의 해결책은 (19)를 만족

9) Pollack[1970], 79 쪽.

하는 술어 $T$를 대상언어 $\mathcal{L}$에서 추방하여 그것의 상위언어로 가져다 놓는 것이다. 그에 따라 그러한 $T$에 대응하는 자기지시적인 문장, 즉 (18)을 만족하는 문장 $\chi$도 자동적으로 존재하지 않게 된다. 하나의 대상언어에 그 상위언어를 인정함으로써 타르스키는 그 상위언어에서 더 이상의 상위언어로 가는 길을 열어놓게 되며 이렇게 해서 언어가 계층화됨과 동시에 진리 개념도 언어에 따라 계층화가 이루어지게 된다. 이에 반해 크립키는 술어 $T$를 대상언어 안에 유지하는 대신에 그 술어가 (19)를 만족시키는 대신

$$\text{(20)} \qquad \mathscr{M} \vDash T(`\phi\text{'}) \Leftrightarrow \mathscr{M} \vDash \phi,$$
$$\mathscr{M} \VDash T(`\phi\text{'}) \Leftrightarrow \mathscr{M} \VDash \phi$$

를 만족시키도록 하는 전략을 취하고 있다. 실제로 (18)과 (20)으로부터 얻어지는 관계는 $\mathscr{M} \vDash T(`\chi\text{'}) \Leftrightarrow \mathscr{M} \VDash T(`\chi\text{'})$와 $\mathscr{M} \vDash \chi \Leftrightarrow \mathscr{M} \VDash \chi$이다. 그러므로 $\chi$와 $T(`\chi\text{'})$ 모두 진리치 간격에 위치하게 된다. 다시 말해 참도 거짓도 아닌 제 3의 진리 값을 갖게 된다.

부동점 $\mathscr{M}$에 속하는 해석 함수 $v$에 대해 $v(T) = \langle E, A \rangle$라고 하자. '$\mathscr{M} \VDash$'를 정의하는 위의 구절 3에 의해 모든 $\phi$에 대해 $\phi \in E \Leftrightarrow \mathscr{M} \vDash T('\phi') \Leftrightarrow \mathscr{M} \VDash \neg T('\phi')$이고 $\phi \in A \Leftrightarrow \mathscr{M} \VDash T('\phi') \Leftrightarrow \mathscr{M} \vDash \neg T('\phi')$. 그러므로 술어 $F$를 $\neg T$로 정의하면 $v(F) = \langle A, E \rangle$라고 말할 수 있다. 이제 $M - A \cup E$를 외연으로 갖고 $A \cup E$를 반외연으로 갖는 술어 $U$를 도입하기로 하자. 즉 $v(U) = \langle M - A \cup E, A \cup E \rangle$. $U(x)$는 곧 $x$가 참도 아니고 거짓도 아님을 의미한다. $U$의 정의에 따라 다음을 쉽게 알 수 있다.

$$\text{(21)} \qquad \mathscr{M} \vDash F('\phi') \vee U(`\phi\text{'}) \Leftrightarrow \mathscr{M} \nvDash T('\phi')$$

다음과 같은 자기지시 문장을 생각해 보자.

$$\text{(22)} \qquad \delta = F('\delta') \vee U(`\delta\text{'})$$

$\delta$는 즉 자신이 거짓이거나 진리 값이 정의되지 않은 문장임을 말하는 자기지시 문장이다. (22)와 위의 (20)에 의해

$$(23) \qquad \mathscr{M}\vDash F('\delta')\vee U(`\delta') \Leftrightarrow \mathscr{M}\vDash T('\delta')$$

그런데 (21)로부터

$$(24) \qquad \mathscr{M}\vDash F('\delta')\vee U(`\delta') \Leftrightarrow \mathscr{M}\nvDash T('\delta')$$

(23)과 (24)로부터 $\mathscr{M}\vDash T('\delta') \Leftrightarrow \mathscr{M}\nvDash T('\delta')$라는 모순된 결론을 얻을 수 있다.

시몬스(Simmons, K.)는 (22)에서와 같은 δ를 대각선 논법을 이용해서 얻고 있는데 그것은 본질적으로 다음과 같은 대입함수를 이용하는 것과 맞먹는 방법이다. 지금 "$z$는 $x$가 가리키는 술어에서 변항 $x$ 대신에 $y$가 가리키는 술어를 바꾸어 놓은 결과 얻어지는 문장을 가리킨다."라는 관계를 생각해 보자. 괴델(K. Gödel)이 보인 것처럼 형식 문장 $\mathscr{L}$이 산수학을 표현할 수 있을 만큼 강력하면 그러한 관계를 표현하는 함수 $\sigma(x, y)=z$가 $\mathscr{L}$에 존재한다. 그러므로 예를 들어 $\sigma(`\phi(x)', '\psi(x)')$는 문장 $\phi('\psi(x)')$를 가리킨다. 즉 $\mathscr{M}\vDash\sigma(`\phi(x)', '\psi(x)')=`\phi('\psi(x)')'$. 여기서 다음과 같은 술어 $\kappa(x)$를 생각해 보자.

$$(25) \qquad \kappa(x) = F(\sigma(x, x))\vee U(\sigma(x, x))$$

(25)에서 $x$ 대신에 '$\kappa$'를 대입하면 다음을 얻는다.

$$(26) \qquad \kappa(`\kappa') = F(\sigma(`\kappa', `\kappa'))\vee U(\sigma(`\kappa', `\kappa'))$$

σ의 정의에 의해,

$$(27) \qquad \mathscr{M}\vDash\sigma(`\kappa', `\kappa')=`\kappa(`\kappa')'$$

따라서 (26)과 (27)로부터

$$(28) \qquad \mathscr{M}\vDash\kappa(`\kappa')\leftrightarrow F(`\kappa(`\kappa')`)\vee U('\kappa(`\kappa')`)$$

(28)은 $\kappa(`\kappa')$가 바로 (22)의 δ와 마찬가지로 자신이 참이거나 진리 값이 정의되지 않았음을 말하는 자기지시 문장임을 보여준다.

위의 논의는 '모델 $\mathscr{M}$에서 참도 거짓도 아니다'라는 의미론적 개념을

의미하는 술어가 $\mathcal{L}$에 존재한다고 할 경우 여전히 모순이 야기됨을 보여주고 있다. 따라서 시몬스는 그러한 의미론적 개념은 '참', '거짓'과는 달리 여전히 대상 언어에서 표현이 불가능하다는 결론을 내리고 있다.[10] 그러한 개념을 표현하기 위해서는 메타언어로 올라갈 필요가 있다. 그러나 그것은 크립키의 진리론이 안고 있는 결정적인 문제점은 아니다.

(20)을 만족하도록 크립키가 복잡한 과정을 거쳐서 고안한 진리 개념은 거짓말쟁이 문장을 참도 거짓도 아닌 진리 값이 없는 문장으로 분류함으로써 '참'의 의미를 다시 설명한 다음, 그러한 설명을 토대로 거짓말쟁이의 문장이 역설을 야기하는 것을 효과적으로 봉쇄하고 있다. 이러한 점에서 역설을 대처하는 방안에 관한 폴락의 지침을 충실히 따르고 있다고 말할 수 있다. 그러나 그것과 크립키의 진리 정의가 어떤 가치를 지니는가 하는 것과는 별개의 문제이다. 여기서 언어 $\mathcal{L}$에서 크립키가 정의한 진리 개념을 $T_1$이라고 하고 (19) 관계를 만족시키는 진리 개념을 $T_2$라고 하자. 그리고 어떤 사람이 "모든 크레타 인이 한 말은 '참'이 아니다."라고 말했다고 하자. 이 경우 그가 언급한 '참'은 (19)를 만족시키는 진리 개념 $T_2$와 크립키가 고도로 복잡한 집합론과 논리 이론을 사용하여 정의한 $T_1$ 가운데 어느 것에 더 가까울까?

우선 일상적인 의미의 '참'의 의미가 $T_1$과 같을 수는 없다. 만일 그러했다면 역설이 야기될 리 없을 것이다. 우리가 위와 같은 사람이 사용한 '참'이 역설을 야기한다고 말하는 이유는 그것의 의미가 $T_2$와 가깝기 때문이었다고 해야 한다. 따라서 크립키의 진리론이 거짓말쟁이의

10) Simmons[1990], 294 쪽. 김도식 교수는 Kim[2004], 80-2 쪽에서 크립키가 이러한 곤경에서 벗어날 수 있음을 논증하고 있는데 그의 논의에는 다소 모호한 점이 있다. 김교수는 시몬스와 다른 방식으로 대각선 위치에 진리치를 채울 경우, 모순을 유발하지 않는 문장이 나온다는 것을 보이고 있는데 중요한 것은 시몬스와 같은 방식으로 대각선 논법을 적용할 경우 모순을 유발하는 문장이 나온다는 점이다. 그리고 시몬스와 같은 방식으로 대각선 논법을 적용하면 (22)에서와 같은 자기 지식적인 문장이 얻어진다는 점이 김도식 교수의 논의에도 불구하고 부정되는 것 같지 않다.

역설을 해결했다고 하면 그것은 그가 정의한 진리 개념이 바로 우리가 사용하고 있는 진리 개념에 대한 보다 정확한 분석이었기 때문은 아니다. 오히려 크립키의 진리 정의보다는 (19)가 일상적인 진리 개념에 대한 정확한 분석이라고 해야 할 것이다. 치하라(Chihara, C.)가 지적한 대로 일상적인 언어를 사용하는 사람들이 크립키가 정의한 것 같은 고도로 복잡하고 논리적으로 세련된 진리 개념을 가지고 있다고 한다면 그렇게 복잡한 개념을 어떻게 획득했는가를 설명해야 하는 지극히 어려운 부담을 안게 될 것이다.[11] 그렇다면 크립키의 진리 정의의 가치는 어디에서 찾을 것인가?

$T_2$와 같은 의미에서의 '참'의 개념은 역설을 야기하므로 그대로 둘 수는 없다. 그렇다면 '참'을 이제부터 $T_1$과 같은 의미로 사용하도록 권장할만한 가치가 있을까? 일상적인 사람들은 $T_1$과 같은 복잡하고도 세련된 진리 개념을 가지고 있지도 않거니와 그러한 개념을 일상적으로 사용하도록 권장할 수는 더더욱 없는 노릇이다. 그러한 개념은 이론적인 관점에서는 아무리 세련되고 심오하다고 할지라도 일상적인 삶을 영위하는 사람들이 배울 것으로 기대하기는 무리다. 따라서 일상인들이 새로 배울 가치가 있다는 것으로 그 개념을 정당화하기는 어렵다. 일반적으로 수학이나 물리학과 같은 학문적 언어는 정밀하게 짜여 있어 거의 모순을 야기할만한 표현이 등장하지 않는다. 그러나 일반인들이 다루기는 어렵다. 반면에 일상어는 다루기 용이한 대신에 '참'과 같은 모순을 야기하는 용어가 등장하기 쉽다. 그러나 설사 다소 모순을 야기하더라도 다루기 쉽다는 것이 일상적인 표현이 지니는 장점이다. 그러한 표현을 수학적인 엄밀한 용어로 대체하는 것은 그러한 표현이 야기하는 모순에 대한 바람직한 해결책이라고 생각할 수 없다. 크립키 자신도 그가 제시하는 모델이 한편으로 형식적 구조와 수학적 성질들이 풍부

11) Chihara[1979], 610 쪽. 크립키 자신도 Kripke[1975], 699 쪽에서 자신이 제안한 진리론이 통상적인 '참'의 용법에 대한 해석을 제시하려는 것이 아니며 의미론적인 역설들에 대한 해결책을 제시하려는 것도 아님을 명시적으로 밝히고 있다.

하게 있는 영역들을 제시하고 있으며 그러한 성질들이 중요한 직관을 상당한 정도로 담고 있다고 말하고 있다.[12] 그러나 형식적 구조와 수학적 성질들이 풍부하다는 것은 이론적으로 흥미를 유발할 수는 있어도 그 모델의 철학적 정당화는 반드시 관련이 있는 것은 아닐 것이다. 크립키의 진리 개념이 중요한 직관을 상당히 담고 있는 것은 사실이지만 그러한 개념이 제시된 고도로 복잡한 수학적 장치는 그와 같은 장점을 상당히 상쇄시키고 있는 것도 사실이다. 크립키의 진리론은 진리 개념의 바람직한 정의는 단순히 진리 개념에 대한 새로운 정의 내지는 설명을 통해 역설이 야기되는 것을 방지하는 것만으로는 부족하며 철학적인 정당성을 동반해야 하는 문제라는 것을 가르치고 있다.



12) Kripke[1975], 같은 곳.

# 참고문헌

Barker, J.(1999), *The Inconsistency Theory of Truth*, 프린스턴(Princeton) 대학 박사 학위 논문.

Barwise, J and Etchemendy, J.(1987), *The Liar*, Oxford University Press.

Chihara, C.(1979), "The Semantic Pradoxes: A diagnostic Investigation", *The Philosophical Review* **88**, 590-618 쪽.

Gupta, A.(1982), "Truth and Paradox", *Journal of Philosophical Logic* 11, 1-60 쪽, Martin(1984), 175-235 쪽에 전재.

Herzberger, H.(1982), "Notes on Naïve Semantics", *Journal of Philosophical Logic* 11, 61-102 쪽, Martin(1984), 133-74 쪽에 전재.

Kim, Doesik(2004), "Kripke's Theory of Truth and the Liar Pradox", 논리연구 제 7집 제 1호, 67-83 쪽.

Kleene, S. C.(1952), *Introduction to Metamathematics*, Elsevier.

Kripke, S. A.(1975), "Outline of a Theory of Truth", *Journal of Philosophy* **72**, 690-716 쪽.

Martin, R.(ed.)(1970), *The Paradox of the Liar*, Ridgeview Publishing Company.

____________(1984), *Recent Essays on Truth and the Liar Paradox*, Oxford University Press.

McGee, V.(1991), *Truth, Vaguness, and Paradox*, Hackett Publishing Company.

Pollack, J.(1970), "The Truth about Truth: A Reply to Brian Skyrms", Martin(ed.)(1970), 79-89 쪽에 게재.

Simmons, K.(1990), "The Diagonal Argument and the Liar", *Journal of Philosophical Logic* 19. 277-303 쪽.

Tarski, A(1935), "Der Warheitsbegriff in den formalisierten Sprachen", *Studia Logica* **1**, 261-405 쪽, J. H. Woodger 번역(1956), “The Concept of Truth in Foramalize Languages", J. Cocoran 편집, *Logic, Semantics, Mathematics*, Hackett Publishing Company, 152-278 쪽에 수록.

_____________(1944), "The Semantic Conception of Truth), *Philosophy and Phenomenological Research* **4**, L. Linsky 편집, *Semantics and Philosophy of Language*, University of California Press, 1972, 13-47 쪽에 전재.

## The Semantical Paradox and the Concept of Truth

Jong-Kwon, Lee

Alfred Tarski and Saul Kripke are among the philosophers who believe that the Liar Paradox presses for modification of the concept of truth. They attempted to 'solve' the paradox by showing how their definitions of truth can dissolve the paradox. In this paper, I investigated how Kripke attempted to overcome the problems that Tarski's definition of truth faces and what kind of difficulties Kripke's theory of truth faces in its turn.

Key Words: Semantical Paradox, Liar Paradox, Theory of Truth, Tarski, Kripke.

이종권 e-mail : leejk@cau.ac.kr