Formale Wahrheitstheorien nach Alfred Tarski

Michael Pucher

4. April 2001

Inhaltsverzeichnis

Kapitel 0

Einleitung

Zuerst etwas zum ambitionierten Titel meiner Arbeit. Es ist unmöglich in einer Diplomarbeit die gesamte Entwicklung der Wahrheitstheorien nach Tarski darzustellen. Ich möchte daher vesuchen die formalen Grundlagen dieser Entwicklung herauszuarbeiten, und dann zwei Proponenten dieser Entwicklung besprechen.

Ich habe versucht Resultate der formalen Logik aus den unterschiedlichsten Gebieten wie Mengetheorie, Modelltheorie usw. zu finden, die mit der Entwicklung formaler Wahrheitstheorien in Verbindung stehen. Die Auswahl ist eine subjektive, und einige Themen habe ich sehr genau bearbeitet, z.b. das Diagonallemma und die Technik der Diagonalisierung andere wiederum sind nur zusammengefasst (z.b. die partiellen Wahrheitsprädikate).

Die Auswahl die ich vorgenommen habe, spiegelt den Wechsel meiner Interessen während ich an dem Thema gearbeitet habe. Zu Beginn wollte ich einige Aufsätze endlich einmal im Original lesen. Das waren der Aufsatz von Georg Cantor [Cantor:1962] über die Mächtigkeit des Kontinuums und der Aufsatz von Alfred Tarski [Tarski:1939] über die Undefinierbarkeit der Wahrheit. Dann hat mich vor allem die Definierbarkeit von Wahrheitsprädikaten für Teilsprachen und die Definierbarkeit in der Logik 2. Stufe interessiert.

Die mathematische Logik nimmt einen so grossen Teil der Arbeit ein, weil ich glaube dass man ohne Kenntnis der mathematischen Logik die von mir behandelten Autoren nicht wirklich verstehen kann. Ausserdem habe ich immer versucht in den Kapiteln über formale Logik philosophische Fragen aufzuwerfen und zu diskutieren. Das so etwas nur anzatzweise möglich ist, wird jedem klar sein der sich schon mit der Philosophie der Logik beschäftigt hat.

Die Autoren, mit denen ich mich genauer beschäftigen möchte sind Saul Kripke und Tyler Burge. Der erste wird Saul Kripke sein, der mit seiner Theorie wichtige Ideen in die Diskussion eingebracht hat. Der zweite wird Tyler Burge sein, der mich interessiert hat weil sein Ansatz über die Tarski-Kripke'sche Tradition hinausgeht.

Obwohl sich ein grosser Teil dieser Arbeit mit den formalen Grundlagen von Wahrheitstheorien beschäftigt, wird die Tarski'sche Wahrheitstheorie nicht explizit bearbeitet werden. Aber immer dort, wo modelltheoretische Begriffe verwendet werden, ist Tarski's Theorie präsent, weil seine Ideen die Basis für die moderne Modelltheorie geliefert haben. Ich habe mich zu diesem Vorgehen entschlossen, um die Notation so einheitlich wie möglich zu halten. Überhaupt war es eine der interessantesten Aufgaben dieser Arbeit, die unterschiedlichen Notationen ineinander zu übersetzen.

Mich hat interessiert warum eine formale Wahrheitstheorie akzeptiert wird. Welche Ergebnisse der formalen Logik müssen berücksichtigt werden ? Eines der zentralen Theoreme in diesem Zusammenhang ist das Theorem von der Undefinierbarkeit der Wahrheit. Es bildet den Hintergrund auf dem viele Theorien verstanden werden können. Warum man sich z.b. mit partiellen Wahrheitsprädikaten beschäftigt.

Das Hauptproblem der hier vorgestellten Theorien besteht in der Definition eines Wahrheitsprädikats. Wäre die Definition eine einfache Sache, dann würde man sich nicht mehr länger dabei aufhalten. Die Definition zeigt sich allerdings als schwierig bis unmöglich.

Definitionen können aus verschiedensten Gründen schwierig sein. Versuchen wir zum Beispiel das Prädikat "Tisch" zu definieren, d.h. notwendige und hinreichende Bedingungen anzugeben die etwas erfüllen muss um ein Tisch zu sein.

Die Definition kann zu weit sein (einige Dinge die keine Tische sind fallen darunter) oder sie kann zu eng sein (einige Tische fallen nicht darunter). Hier könnte man sich fragen, welchen Sinn eine Definition überhaupt haben kann, wenn man doch schon wissen muss was ein Tisch ist um die Definition verstehen und korrigieren zu können. Die Anwort lautet "Begriffsklärung". Die Definition hilft uns genauer zu verstehen was wir unter einem Tisch verstehen. Das implizite Wissen wird explizit gemacht.[1]

Die Definition kann unmittelbar oder mittelbar zirkulär sein. Zur Defini-

[1]In der Formalwissenschaft haben Definitionen aber auch eine andere Verwendung. Dort werden aus einfachen Begriffen die man schon kennt komplexere Begriffe definiert.

tion des Begriffs wird der Begriff selbst unmittelbar oder über die Definition eines anderen Begriffs verwendet. Die Zirkularität einer Definition stellt ein Problem dar, wenn wir annehmen dass die Definition in gewisser Weise das Erlernen eines Begriffs wiederspiegelt. Denn in diesem Fall müssen wir uns fragen wie es möglich war und ist die Verwendung dieses Begriffs zu lernen. Wenn wir lernen als einen hierachischen Vorgang verstehen dann haben wir ein Problem.

Aber selbst wenn wir all das nicht anerkennen, werden wir akzeptieren dass eine nichtzirkuläre Definition einer zirkulären Definition vorzuziehen ist. Die nichtzirkuläre Definition ist einfach informativer. Andernfalls könnten wir behaupten alle definitorischen Probleme ein für alle Mal gelöst zu haben. Jeder Begriff ist durch sich selbst definiert. Die zirkuläre Definition ist der Nullpunkt der Definition.

Eine zirkuläre Definition wird wahrscheinlich nur dann akzeptiert werden, wenn keine andere Definition gegeben werden kann. Das keine andere Definition gegeben werden kann muss zusätzlich zur Angabe der zirkulären Definition bewiesen werden.

Bei der Definition semantischer Begriffe wie "Wahrheit" stossen wir allerdings auf ein spezielles Problem. Beim Verständnis dessen was eine Definition ist, spielen semantische Begriffe nämlich eine wesentliche Rolle. Es scheint so als müssten wir bei diesen Definitionen die zu definierenden Begriffe immer schon voraussetzen. Aus diesem Grund hat Gottlob Frege "Wahrheit" für undefinierbar gehalten.

Ein anderes Problem ist die Möglichkeit der Selbstanwendung. Wenn man einen semantischen Begriff definiert, dann gehört das definierende Prädikat in die Kategorie der Gegenstände auf die es angewendet werden kann. Diese Selbstanwendung ist nicht immer problematisch, kann aber bei der Definition von "gewöhnlichen"Begriffen nicht auftreten.

Es wird sich zeigen, dass diese Schwierigkeiten im Falle des Wahrheitsprädikats tatsächlich zum Problem werden. Tarski hat bewiesen, dass es nicht möglich ist in einer hinreichend interessanten Sprache ein totales Wahrheitsprädikat für diese Sprache zu definieren. Der Beweis folgt nicht den eben gemachten Überlegungen, sondern geht aus von bestimmten selbstreferentiellen Sätzen, die in der Sprache formuliert werden können, aber nicht bewertet werden können. In diesem Zusammenhang ist es notwendig einige Paradoxien genauer zu untersuchen.

0.1 Paradoxien

Warum stellt eine Paradoxie ein Problem dar? Für den Logiker ist die Antwort wahrscheinlich klar. Eine Paradoxie stellt ein Problem dar, weil aus ihr alles folgt. In der klassischen Logik gibt es die Regel *ex falso quod libet* (aus falschem beliebiges) nach der aus einer widersprüchlichen Aussage jede beliebige Aussage abgeleitet werden kann.

Kann aber eine widersprüchliche Aussage abgeleitet werden, die von keiner anderen Aussage abhängt (handelt es sich also um eine Paradoxie) dann kann jede beliebige Aussage bewiesen werden. Die Paradoxie ist damit nicht nur ein Problem einer speziellen Theorie, sondern ein Problem für Theorien im Allgemeinen. Wird eine Paradoxie akzeptiert, dann kann jede Theorie *ad absurdum* geführt werden.

Nimmt man an, dass es eine notwendige Eigenschaft von Theorien ist[2], die Sätze die in der Sprache der Theorie formuliert werden können in wahre und falsche Sätze aufzuteilen, so folgt aus der Akzeptanz einer Paradoxie dass es nichts gibt was diese Eigenschaft haben könnte, dass es also keine Theorien gibt.

Ein zusätzliches Problem ist, dass wir die Folgen einer Paradoxie gar nicht auf diese Art und Weise formulieren können. Die Schlüsse die wir gezogen haben, sind nur dann möglich wenn wir uns auf einer Metaebene bewegen, die Paradoxie also nicht ernst nehmen oder als etwas ansehen das uns nichts angeht. Ansonsten müssten wir akzeptieren, dass alle Schlüsse die wir gezogen haben in irgendeiner Art und Weise auf einer Unterscheidung zwischen wahren und falschen Sätzen beruhen, diese Unterscheidung aber gerade durch die Paradoxie in Frage gestellt wird.

Die Schwierigkeit einen Standpunkt zu formulieren der die Paradoxie akzeptiert, kann als Hinweis gesehen werden dass es einen solchen Standpunkt nicht gibt.

Eine Möglichkeit mit diesem Problem umzugehen ist die Entwicklung einer mehrwertigen Logik. Das Problem besteht allerdings weiter als das Problem eine mehrwertige Logik konsequent auch für die Metasprache zu entwickeln. In der mehrwertigen Logik kann ein Satz nicht nur wahr oder

[2]D.h. Wenn etwas eine Theorie ist, dann hat dieses etwas diese Eigenschaft. Es ist nicht notwendig, dass die Theorie für jeden Satz angibt ob er wahr oder falsch ist. Wenn laut Theorie einige Sätze wahr sind ist die Unterscheidung schon gemacht, weil deren Negation dann falsch ist

falsch sein, sondern beliebig viele andere Wahrheitswerte haben. Die Metasprache der mehrwertigen Logik ist dabei allerdings die klassische zweiwertige Logik. Ein Satz ist entweder wahr oder falsch, oder hat einen der beliebigen anderen Wahrheitswerte. Will man sich klarmachen, was *mehrwertige Metasprache* bedeutet, dann muss man in die Metametasprache wechseln, die wiederum klassisch ist.

Nehmen wir an wir hätten eine mehrwertige Logik mit drei Wahrheitswerten *wahr, falsch* und *unbestimmt*, und eine mehrwertige Metasprache mit denselben Wahrheitswerten. Dann könnten wir in einer klassischen Metametasprache Sätze bewerten wie:

Die Unbestimmtheit von **p** *ist unbestimmt*, *Die Wahrheit von* **p** *ist unbestimmt*, *Die Falschheit von* **p** *ist unbestimmt* usw.

wobei diese Sätze entweder wahr oder falsch sind. Die Schwierigkeit eine konsequente mehrwertige Logik zu entwickeln spielt meiner Meinung nach auch beim Problem des *verstärkten Lügners* eine Rolle, dass ich im Abschnitt über Tyler Burge besprechen werde.

Nachdem nun einiges über Paradoxien im Allgemeinen gesagt wurde, möchte ich etwas über spezielle Paradoxien sagen.

Die sogenannte Lügnerparadoxie hat in der Entwicklung der formalen Wahrheitstheorien eine besondere Rolle gespielt. Viele Wahrheitstheorien sind formuliert worden um diese Paradoxie zu lösen. Auch Saul Kripke nimmt dieses Paradoxon als einen Ausgangspunkt und seine Lösung als Erfolg seiner Theorie.

Dieser Satz ist falsch

Versucht man den Satz zu bewerten, kommt man zu dem Schluss dass er sowohl wahr als auch falsch sein muss. Das aber ist unmöglich also gibt es ein Problem mit dem Satz oder mit der Bewertung. Auf den ersten Blick sieht der Satz etwas ungewöhnlich aus, weil er so gar nichts interessantes aussagt. Eben lediglich etwas über sich selbst. Wie zum Beispiel auch:

Dieser Satz besteht aus sechs Wörtern

Eine einfache Lösung der Paradoxie bestünde darin diesen Ausdruck nicht als Satz zu verstehen und so von der Bewertung auszuschliessen. Vielleicht auch mit der Hoffnung, dass solche Sätze in einer syntaktisch und semantisch exakt definierten Sprache nicht vorkommen. Die Entwicklungen der modernen Logik haben diese Hoffnung allerdings zunichte gemacht. Der Begriff des selbstreferentiellen Satzes hat eine exakte Definition gefunden. Der Satz kann also nicht aufgrund seiner Nichtsatzartigkeit von der Bewertung ausgeschlossen werden.

Ein anderer Ausweg bestünde darin die Irrelevanz des Satzes zu betonen. Ein Satz der nur etwas über sich selbst sagt trägt zum besseren Verständnis der Welt nicht sehr viel bei. Diese Lösung scheint vielleicht am akzeptabelsten zu sein, ist aber von einem logischen Standpunkt aus eher unbefriedigend. Der Begriff der Irrelevanz ist eben ein relativer, und was für die eine Wissenschaft irrelevant ist, ist für eine andere vom zentralem Interresse.

Ein anderes Paradoxon, dass für meine Arbeit von Interesse ist, ist *Richard's Paradox*. Es ist deshalb von Interesse, weil eine ähnliche Konstruktion im Gödelschen Beweis verwendet wird.

Die Begriffswörter[3] können in *homologische* und *heterologische* unterteilt werden. Ein Begriffswort ist *homologisch*, wenn das Begriffswort unter den Begriff fällt den es ausdrückt[4], andernfalls ist es *heterologisch*.

Jetzt kann man sich fragen, ob "*heterologisch*" *heterologisch* ist oder nicht.

Wenn es *heterologisch* ist, dann fällt es unter den Begriff den es ausdrčkt ist also *homologisch*. Wenn es nicht *heterologisch* ist, dann fällt es nicht unter den Begriff den es ausdrückt, ist also nicht *homologisch*, also ist es *heterologisch*.

"*Heterologisch*" ist also *heterologisch* genau dann wenn "*heterologisch*" nicht *heterologisch* ist.

Die Verbindung zum Gödelschen Beweis wird klar, wenn man die Sätze einer Sprache bezüglich eines formalen Systems in beweisbare und nicht beweisbare teilt (analog der Unterscheidung *homologisch* und *heterologisch*), und sich dann fragt ob der Satz der von sich selbst behauptet nicht beweisbar zu sein beweisbar ist. Diese paradoxe Struktur[5] kann man verwenden um zu zeigen, dass weder dieser Satz noch seine Negation bewiesen werden können.

In diesem Fall ist der Satz aber wahr. Es wurde ein wahrer Satz gefunden, der weder beweisbar noch widerlegbar ist. Das untersuchte System ist also unvollständig. Nicht alle wahren Sätze können bewiesen werden.

Da man vor diesem Beweis angenommen oder gehofft hatte, dass "einfa-

[3]Wörter die an einen singulären Ausdruck angehängt einen Satz ergeben.

[4]z.b. "Hat drei Wörter" hat drei Wörter.

[5]Der Satz der von sich behauptet nicht beweisbar zu sein, und die Frage ob er beweisbar ist oder nicht.

che" mathematische Systeme wie die Theorie der natürlichen Zahlen vollständig seien war man von diesem Beweis überrascht bzw. geschockt.

Soviel zu meiner Motivation diese Arbeit zu schreiben und zu den historischen Hintergründen dieser Arbeit.

Kapitel 1

Logische Grundlagen

In diesem Kapitel möchte ich die modelltheoretischen ,rekursionstheoretischen und mengentheoretischen Begrife erarbeiten, die ich in den darauffolgenden Kapiteln verwenden werde. In der Modelltheorie wird, im Anschluss an Tarski, der Begriff der Wahrheit mittels des Begriffs der Erfüllbarkeit definiert.

Wie die Resultate der folgenden Kapitel zeigen werden, ist es für alle interessanten Theorien bzw. Sprachen unmöglich, den Begriff der Wahrheit in diesen Theorien selbst zu definieren. Darum wird der Begriff eines wahren Satzes einer Objektsprache in einer Metasprache definiert. Diese Methode setzt immer schon den Wahrheitsbegriff der Metasprache voraus, der in einer Metasprache für die Metasprache explizit gemacht werden kann.

1.1 Modelltheoretische Grundlagen

Alle formalen Sprachen die ich untersuchen werde sind Sprachen der Prädikatenlogik 1.Stufe.[1] Um eine Sprache anzugeben genügt es die Menge von nichtlogischen Symbolen der Sprache festzulegen.

Die logischen Symbole der Prädikatenlogik erster Stufe mit Identität sind: Die Konnektive $\wedge$ (und),$\vee$ (oder), $\neg$ (nicht), $\rightarrow$ (wenn dann), $\leftrightarrow$ (genau dann wenn), das Gleichheitssymbol $=$, die Quantoren $\exists$(es gibt), $\forall$(für alle), Klammernsymbole (,) und abzählbar unendlich viele Variablen $x, y, z, x_1, \ldots$. Ich würde auch mit den Symbolen $\vee$, $\neg$ und $\exists$ auskommen und könnte alle

[1]Bis auf eine Ausnahme im Kapitel über Definierbarkeit in der Logik zweiter Stufe.

anderen logischen Symbole definieren. Diese Methode wird bei beweistheoretischen Untersuchungen verwendet, da Induktionsbeweise meistens nach der Komplexität der Formeln geführt werden. Je weniger komplex die Formeln sind, desto einfacher die Induktionsbeweise.

Eine Sprache L ist eine Menge von Prädikatsymbolen, Funktionssymbolen und Konstantensymbolen. Die Terme, Formeln und Sätze von L werden induktiv definiert.

Definition 1.1.1 *Die Terme von L sind die kleinste Menge von Zeichenketten welche die Variablen und alle Konstantensymbole von L enthalten und unter folgender Formationsregel abgeschlossen sind: Wenn $t_1, ..., t_n$ Terme von L sind und f ein n-stelliges Funktionssymbol von L ist, dann ist $f(t_1, .., t_n)$ ein Term von L.*

Die Formulierung "kleinste Menge" garantiert, dass nur Zeichenketten in der Menge der Terme enthalten sind die unsere Bedingungen erfüllen.

Definition 1.1.2 *Eine Atomformel von L ist entweder von der Form $t_1 = t_2$, wobei t_1 und t_2 Terme von L sind, oder von der Form $R(t_1 \dots t_n)$, wobei R ein n-stelliges Prädikat von L ist und $t_1 \dots t_n$ Terme von L sind.*

Definition 1.1.3 *Die Formeln von L sind die kleinste Menge von Zeichenketten, welche die Atomformeln enthalten und unter folgenden Formationsregeln abgeschlossen sind:*

(i) Wenn ϕ und ψ Formeln sind, dann sind auch

$$\neg\phi, (\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$$

Formeln von L.

(ii) Wenn ϕ eine Formel ist und v eine Variable ist, dann sind $(\exists v\phi)$ und $(\forall v\phi)$ Formeln.

Definition 1.1.4 *Die Menge der freien Variablen einer Formel ϕ, $FV(\phi)$ ist folgendermassen definiert:*

(i) Wenn ϕ eine Atomformel ist, dann ist $FV(\phi)$ gleich der Menge von Variablen die in ϕ vorkommen.

(ii) $FV(\neg\phi) = FV(\phi)$,

(iii) $FV(\phi \wedge \psi) = FV(\phi \vee \psi) = FV(\phi \rightarrow \psi) = FV(\phi \leftrightarrow \psi) = FV(\phi) \cup FV(\psi)$,

(vi) $FV(\exists v\phi) = FV(\forall v\phi) = FV(\phi) - \{v\}$

Definition 1.1.5 *Ein Satz von L ist eine Formel von L ohne freie Variablen.*

Weiters muss noch eine Erfüllungsrelation $\mathfrak{M} \models \phi$ zwischen mengentheoretischen Strukturen $\mathfrak{M}$ und Sätzen der Sprache L definiert werden. Während die Definition der Sätze für die gesamte Untersuchung beibehalten werden kann, wird es notwendig sein verschiedene Arten von Strukturen zu definieren.

$\mathfrak{M}$ ist eine Struktur für eine Sprache L, wenn $\mathfrak{M} = (D, \Phi, I)$ ist, wobei D eine nichtleere Menge von Objekten ist(das Diskursuniversum), Φ eine Funktion ist, die jeder Konstante c von L ein Element $\Phi(c)$ aus D, jedem n-stelligen Prädikatbuchstaben P aus L eine Teilmenge $\Phi(P)$ von D^n und jedem n-stelligen Funktionssymbol f aus L eine Funktion $\Phi(f)$ vom Typ $D^n \rightarrow D$ zuordnet.I ist eine Funktion(die Variablenumgebung) die jeder Variable v ein Element $I(v)$ aus D zuordnet. Der Referent eines Terms ergibt sich aud den Referenten der Namen und den Referenten der Variablen.

Definition 1.1.6 *Sei $\mathfrak{M}$ eine Struktur für L. Die Funktion $t^{\mathfrak{M}}$ die den Termen ihre Referenten unter $\mathfrak{M}$ zuordnet ist folgendermassen definiert:*
(i) $t^{\mathfrak{M}}(v) = I(v)$ für alle Variablen v.
(ii) $t^{\mathfrak{M}}(c) = \Phi(c)$ für alle Konstantensymbole c.
(iii) $t^{\mathfrak{M}}(f(t_1, \ldots, t_n) = \Phi(f)(t^{\mathfrak{M}}(t_1), \ldots, t^{\mathfrak{M}}(t_n))$

Aufbauend auf der Definition der Referenten der Terme kann jetzt definiert werden, was es heisst dass eine Struktur $\mathfrak{M}$ einen Satz ϕ erfüllt $\mathfrak{M} \models \phi$. In diesem Fall sagt man auch $\mathfrak{M}$ sei ein Modell von ϕ.

Definition 1.1.7 *Seien $\mathfrak{N}$, $\mathfrak{M}$ Strukturen:*
$\mathfrak{N} \backsim_v \mathfrak{M}$ genaudann wenn sich die beiden Strukturen nur an dem Wert unterscheiden, den ihre Variablenumgebungen der Variable v zuweisen.

Definition 1.1.8 *Sei $\mathfrak{M}$ eine Struktur für L, ϕ ein Satz von L.*
(i) $\mathfrak{M} \models (t_1 = t_2)$ genaudann wenn $t^{\mathfrak{M}}(t_1) = t^{\mathfrak{M}}(t_2)$.
(ii) $\mathfrak{M} \models P(t_1, \ldots, t_n)$ genaudann wenn $(t^{\mathfrak{M}}(t_1), \ldots, t^{\mathfrak{M}}(t_n)) \in \Phi(P)$.
(iii) $\mathfrak{M} \models \neg\phi$ genaudann wenn nicht $\mathfrak{M} \models \phi$.
(iv) $\mathfrak{M} \models (\phi \wedge \psi)$ genaudann wenn $\mathfrak{M} \models \phi$ und $\mathfrak{M} \models \psi$.
(v) $\mathfrak{M} \models (\phi \vee \psi)$ genaudann wenn $\mathfrak{M} \models \phi$ oder $\mathfrak{M} \models \psi$ oder beides.
(vi) $\mathfrak{M} \models (\phi \rightarrow \psi)$ genaudann wenn $\mathfrak{M} \models \phi$ nicht, oder $\mathfrak{M} \models \psi$.
(vii) $\mathfrak{M} \models (\phi \leftrightarrow \psi)$ genaudann wenn $\mathfrak{M} \models \phi$ nicht und $\mathfrak{M} \models \psi$ nicht, oder $\mathfrak{M} \models \phi$ und $\mathfrak{M} \models \psi$.
(viii) $\mathfrak{M} \models (\exists v\phi)$ genaudann wenn es eine Strukur $\mathfrak{N}$ gibt, $\mathfrak{N} \backsim_v \mathfrak{M}$, sodass $\mathfrak{N} \models \phi$.

(ix) $\mathfrak{M} \models (\forall v \phi)$ *genaudann wenn für alle Strukuren* $\mathfrak{N}$, $\mathfrak{N} \backsim_v \mathfrak{M}$ *gilt dass* $\mathfrak{N} \models \phi$.

Bis auf die Behandlung der Variablenbelegungen habe ich alle Definitionen von [Barwise:1977] übernommen. In *(i)* wurde beide Male dasselbe Symbol $=$ verwendet. Der Unterschied in der Bedeutung liegt darin, dass bei der ersten Verwendung die Identitätsrelation der Objektsprache, bei der zweiten die der Metasprache gemeint ist. Definition 2.1.8 kann auch verwendet werden um Formeln mit freien Variablen auszuwerten. Ich werde sie aber nur auf Sätze anwenden. In diesem Fall ist die Variablenumgebung einer Struktur irrelevant.

Wenn eine Struktur einen Satz erfüllt, so ist sie ein Modell des Satzes. Eine Struktur ist ein Modell einer Satzmenge, wenn sie Modell jedes Satzes der Menge ist. Wollen wir sagen, dass jede Struktur Modell eines Satzes ϕ ist so schreiben wir $\models \phi$. In diesem Fall ist ϕ ein gültiger Satz. Sei Γ eine Menge von Sätzen. $\Gamma \models \phi$ bedeutet, dass jedes Modell von Γ auch ein Modell von ϕ ist. In diesem Fall sagen wir das ϕ aus Γ folgt.

$\vdash_K \phi$ bedeutet, dass ϕ in einem logischen Kalkül K ableitbar ist. Wir werden einen Kalkül des natürlichen Schliessens verwenden [Lemmon:1969]. $\Gamma \vdash_K \phi$ bedeutet, dass ϕ aus der Satzmenge Γ im Kalkül K ableitbar ist. Das Subskript K kann wegfallen, wenn klar ist in welchem Kalkül die Ableitung erfolgt.

1.2 Rekursionstheoretische Grundlagen

In diesem Abschnitt werde ich die Klasse der partiell rekursiven Funktionen definieren. Ich möchte nur Funktionen definieren, die im intuitiven Sinne berechenbar sind. Darum beginne ich mit einigen Basisfunktionen die sicher berechenbar sind und gebe Regeln an wie aus diesen Basisfunktionen andere berechenbare Funktionen erzeugt werden können. Jede solcherart erzeugte Funktion nenne ich "partiell rekursiv". Damit ist sichergestellt, dass alle partiell rekursiven Funktionen berechenbar sind. Die Umkehrung dieses Satzes, die Churchsche These, die behauptet dass alle berechenbaren Funktionen partiell rekursiv sind lässt sich nicht formal beweisen. Eine Evidenz für die Churchsche These ist die Tatsache, dass verschiedene Formalismen zur Definition von Funktionen alle dieselbe Funktionenklasse ergeben haben. Es ist zu vermuten, dass diese Klasse mit der Menge der intuitiv berechenbaren

Funktionen zusammenfällt. Die Churchsche These lässt sich nicht formal beweisen, weil es sich um eine erkenntnistheoretische These handelt, die etwas über unsere Möglichkeit Formalismen zu erfinden aussagt. Sie behauptet, dass wir keinen Formalismus erfinden können, mit dem wir mehr als die partiell rekursiven Funktionen definieren können.

Definition 1.2.1 Basisfunktionen
(i) Alle k-stelligen Nullfunktionen $Null_k$ ($k >= 0$)

$$Null_k : (x_1, \ldots, x_k) \to 0$$

sind partiell rekursiv. Für jedes k-Tupel liefert diese Funktion den Wert 0.
(ii) Alle k-stelligen Projektionsfunktionen $Proj_{k,i}$ ($1 <= i <= k$)

$$Proj_{k,i} : (x_1, \ldots, x_k) \to x_i$$

sind partiell rekursiv.
(iii) Die Nachfolgerfunktion N

$$N : x \to x + 1$$

ist partiell rekursiv.

Die in Definition 1.2.1 angegebenen Funktionen sind alle intuitiv berechenbar. Aus partiell rekursiven Funktionen kann mit mittels folgender Regeln neue partiell rekursive Funktionen erzeugen.

Definition 1.2.2 Normierte Einsetzung
Sei Ψ eine partiell rekursive r-stellige Funktion ($r >= 1$) und $\chi_1, \ldots, \chi r$ seien partiell rekursive k-stellige Funktionen, dann ist folgende Funktion Φ eine partiell rekursive k-stellige Funktion:

$$\Phi(x_1, \ldots, x_k) = \Psi(\chi_1(x_1, \ldots, x_k), \ldots, \chi_r(x_1, \ldots, x_k))$$

Definition 1.2.3 Primitive Rekursion *Sei Ψ eine totale k-stellige partiell rekursive Funktion und χ eine totale $k+2$-stellige partiell rekursive Funktion. Dann ist folgende Funktion Φ eine totale $k+1$-stellige partiell rekursive Funktion.*

$$\Phi(x_1, \ldots, x_k, 0) = \Psi(x_1, \ldots, x_k)$$
$$\Phi(x_1, \ldots, x_k, N(y)) = \chi(x_1, \ldots, x_k, y, \Phi(x_1, \ldots, x_k, y))$$

Aus einer Berechnungsvorschrift für Ψ und χ lässt sich einfach eine Berechnungsvorschrift für Φ erstellen. D.h. wenn Ψ und χ berechenbar sind, dann ist auch Φ berechenbar.

Definition 1.2.4 μ-Operator *Sei Ψ eine $k+1$-stellige partiell rekursive Funktion. Dann ist $\mu(\Psi)$ eine k-stellige partiell rekursive Funktion, sodass:*

$$\mu(\Psi)(x_1, \ldots, x_k) = Das\ kleinste\ y\ sodass\ \Psi(x_1, \ldots, x_k, y) = 0$$

$$\wedge \forall z <= y \Psi(x_1, \ldots, x_k, y)\ ist\ definiert.$$

Definition 1.2.5 Partiell rekursive Funktionen

Die Klasse der partiell rekursiven Funktionen ist die kleinste Klasse, die die Basisfunktionen enthält und bezüglich μ-Operator, normierter Einsetzung und primitiver Rekursion abgeschlossen ist.

Definition 1.2.6 Primitiv rekursive Funktionen

Die Klasse der primitiv rekursiven Funktionen ist die kleinste Klasse, die die Basisfunktionen enthält und bezüglich normierter Einsetzung und primitiver Rekursion abgeschlossen ist.

Definition 1.2.7 Primitv rekursive Relationen

Eine Relation wird primitv rekursiv genannt, wenn die charakteristische Funktion der Relation primitiv rekursiv ist.

Wie man leicht sehen kann, sind alle primitiv rekursiven Funktionen total, während es partiell rekursive Funktionen gibt die (wie der Name schon sagt) partiell sind. Durch die Anwendung des μ-Operators kann aus totalen Funktionen eine partielle Funktion entstehen, weil es keinen Wert für y geben muss, der die angegebene Bedingung erfüllt.

1.3 Mengentheoretiche Grundlagen

Zum besseren Verständnis der Kripkeschen Wahrheitstheorie werde ich einige mengentheoretische Begriffe herausarbeiten. Die Extension des Wahrheitsprädikats wird von Kripke als eine Menge verstanden, die durch wiederholte Anwendung eines Operators auf eine andere Menge entsteht. Liefert diese Anwendung keine neuen Elemente mehr, so hat der Operator auf der

Grundmenge einen Fixpunkt. Es ist also wichtig, die Existenz von solchen Fixpunkten zu beweisen.

Die Operatoren haben darüber hinaus die Eigenschaft monoton steigend zu sein, d.h. die aus einer Anwendung des Operators resultierende Menge hat nie weniger Elemente als die Ausgangsmenge.

Die Definitionen und Beweise sind einem Aufsatz von Melvin Fitting entnommen. [Fitting:1986]

Es soll die Existenz von *kleinsten Fixpunkten* und *grössten intrinsischen Fixpunkten* bewiesen werden. Fitting beweist die Sätze die für Kripkes Konstruktion relevant sind in einem mengetheoretischen Kontext, während Kripke in einem modelltheoretischen Kontext arbeitet. Kripke beweist die Existenz von Fixpunkten mittels transfiniter Induktion. Bei einer transfiniten Induktion muss zusätzlich bewiesen werden, dass der Satz auch für die Grenzwerte gilt. Fitting verwendet auch eine andere Beweismethode.

Ich werde zuerst die relevanten Sätze mittels transfiniter Induktion beweisen, und dann die andere Möglichkeit skizzieren. Aus dem einfachen Grund, weil das einfacher ist.

Gesucht wird der kleinste Fixpunkt eines monotonen Operators.

Definition 1.3.1 Halbordung *Eine Menge* $<\mathbf{D}, <=>$ *ist eine Halbordung, wenn* $<=$ *eine Relation auf* $\mathbf{D}$ *ist, die reflexiv, antisymetrisch und transitiv ist.*

Definition 1.3.2 Monoton steigende Operatoren *Ein Operator* Φ *ist monoton steigend auf einer Halbordung* $<\mathbf{D}, <=>$*, wenn für alle Elemente* d, e *aus* $\mathbf{D}$ *gilt, dass*

$$d <= e \rightarrow \Phi(d) <= \Phi(e).$$

Definition 1.3.3 Kleinste obere Schranke a *ist eine obere Schranke von* A*, wenn für alle Elemente* b *aus* A *gilt, dass* $b <= a$. a *ist die kleinste obere Schranke von* A*,* $\sup A$ *wenn alle oberen Schranken von* A *kleiner gleich* a *sind.*

Definition 1.3.4 Transfinite Sequenz *Sei* a *ein beliebiges Element aus* $\mathbf{D}$ *für das* $a <= \Phi(a)$*. Die transfinite Sequenz* a_i *wird folgendermassen definiert.*

$$a_0 = a$$

$$a_{\alpha+1} = \Phi(a_\alpha)$$

$$a_\lambda = \sup\{a_\alpha | \alpha < \lambda\}$$

für alle Grenzwerte λ. λ ist ein Grenzwert, wenn für alle α kleiner als λ gilt, dass auch $\alpha + 1$ kleiner als λ ist. Ein Grenzwert kann also nicht mit der Nachfolgerfunktion von einem kleinerem Element aus erreicht werden.

Diese Sequenz ist steigend, d.h. $a_\alpha <= a_{\alpha+1}$. Für a_0 und $a_{\alpha+1}$ ist das leicht einzusehen, für die Grenzwerte möchte ich es beweisen.

Theorem 1.3.1 *Für die Grenzwerte muss gezeigt werden, dass:*

$$a_\lambda <= a_{\lambda+1}$$

Beweis 1.3.1 *Angenommen λ ist ein Grenzwert, und für alle $\alpha < \lambda$ gilt, dass $a_\alpha <= a_{\alpha+1}$. Nach der Definition von a_λ ist $a_\alpha <= a_\lambda$, also ist $\Phi(a_\alpha) <= \Phi(a_\lambda)$ oder $a_{\alpha+1} <= a_{\lambda+1}$. Mittels Induktionshypothese folgt, dass $a_\alpha <= a_{\lambda+1}$. Weil α beliebig war, ist $\sup\{a_\alpha | \alpha < \lambda\} <= a_{\lambda+1}$ oder $a_\lambda <= a_{\lambda+1}$.*

Es ist leicht einzusehen, dass in dieser Sequenz nicht jedes Element a_i von jedem anderen Element in der Sequenz unterschieden sein kann. Ansonsten wäre die Anzahl der Ordinalzahlen gleich der Mächtigkeit der Menge **D**. Das ist aber unmöglich, weil die Ordinalzahlen keine Menge bilden. Also muss es ein α geben, sodass

$$a_\alpha = \Phi(a_\alpha).$$

Dieses a_α ist ein Fixpunkt der Sequenz. Wenn wir mit dem kleinsten Element aus **D** beginnen, dann erreichen wir einen Fixpunkt, und dieser Fixpunkt ist auch der kleinste Fixpunkt der Sequenz, weil die Elemente ab diesem Fixpunkt nicht mehr grösser werden. Wir haben also bewiesen, dass jede solche Sequenz einen kleinsten Fixpunkt besitzt.

Definition 1.3.5 Intrinsische Fixpunkte *Zwei Fixpunkte sind* **kompatibel**, *wenn sie eine gemeinsame obere Schranke besitzen. Ein Fixpunkt wird* **intrinsisch** *genannt, wenn er mit jedem anderen Fixpunkt* **kompatibel** *ist.*

Solche intrinsischen Fixpunkte können konstruiert werden, wenn man eine abnehmende Sequenz definiert, analog zur Definition einer steigenden Sequenz. Dabei muss man mit einem Element b beginnen für das $b >= \Phi(b)$ gilt. Bei den Grenzwerten verwendet man nicht die *kleinste obere Schranke*, sondern die *grösste untere Schranke*. Beginnt man bei einem Element

b, dann erreicht man mit dieser Sequenz den grössten Fixpunkt unterhalb von b. Den *grössten intrinsischen Fixpunkt* erhält man, wenn die Sequenz mit der grössten unteren Schranke von $\mathbf{M}$ beginnt, wobei $\mathbf{M}$ die Menge der maximalen Fixpunkte von Φ ist.

Jetzt möchte ich noch zeigen, wie man diese Beweise ohne transfinite Induktion führen kann.

Theorem 1.3.2 *Sei* $< \mathbf{D}, <=>$ *eine Halbordnung, wobei* $\mathbf{D}$ *ein kleinstes Element besitzt, jede Kette von* $\mathbf{D}$ *eine obere Schranke besitzt, und jede nichtleere Menge aus* $\mathbf{D}$ *mit einer oberen Schranke eine kleinste obere Schranke besitzt. Ausserdem sei* Φ *ein monotoner Operator auf* $\mathbf{D}$.

Dann gilt:

(i) Wenn $a <= \Phi(a)$, *dann gibt es einen* **maximalen Fixpunkt** *von* Φ *oberhalb von* a *(der grösser ist als* a*).*

(ii) Wenn $a <= \Phi(a)$, *dann gibt es einen* **kleinsten Fixpunkt** *von* Φ *oberhalb von* a *(der grösser ist als* a*).*

(iii) Φ *hat einen* **kleinsten Fixpunkt** .

(iv) Φ *hat einen* **grössten intrinsischen Fixpunkt**.

Ich werde die Beweise für Punkt *(ii)-(iii)* angeben. Der Beweis von *(iv)* und *(iii)* kann in Melvin Fittings [Fitting:1986] Aufsatz nachgelesen werden.

Beweis 1.3.2 *(ii) Angenommen* $a <= \Phi(a)$. *Nach (i) gibt es einen maximalen Fixpunkt von* Φ *oberhalb von* a. *Sei* b *ein solcher Fixpunkt. Dann ist* $a <= b$ *und* $\Phi(b) <= b$. *Mit dem Fixpunkttheorem hat* Φ *eine kleinsten Fixpunkt oberhalb von* a.

(iii) Man nehme das kleinste Element aus $\mathbf{D}$ *in (ii).*

Obwohl also Kripke in seinem Aufsatz, zu dessen Klärung dieser ganze Abschnitt geschrieben wurde, transfinite Induktion verwendet, ist das für diese Theoreme nicht unbedingt notwendig. Das zeigt, dass die verwendeten Strukturen nicht so komplex sind, wie man zuerst annehmen würde. Im Beweis von (i) wird allerdings das Zorn'sche Lemma verwendet, das eines der umstritteneren Axiome der Mengentheorie ist (Da es äquivalent zum Auswahlaxiom ist) . Für den Beweis der beiden Theoreme wird also auch in der zweiten Variante der gesamte mengentheoretische Apparat verwendet.

Kapitel 2

Undefinierbarkeit der Wahrheit

2.1 Diagonalisierung

Die Methode der Diagonalisierung wird im Bereich der Rekursionstheorie und der Definitionstheorie oft angewendet. Ich werde zwei Theoreme besprechen bei deren Beweis die Methode auf zwei unterschiedliche Weisen verwendet wird.

Beim Satz von Cantor wird die Diagonalisierung in einem indirekten Beweis verwendet. Es wird gezeigt, dass die Annahme dass eine bestimmte Abbildung zwischen zwei Mengen existiert, zu einem Widerspruch führt.

Beim Diagonalllemma wird die Methode “konstruktiv” verwendet. Es wird gezeigt, wie aus einem bestimmten Prädikat einer Sprache L in einem hinreichend starken formalen System **S** ein Satz konnstruiert werden kann, der von sich “behauptet” dieses Prädikat zu erfüllen.

2.1.1 Der Satz von Cantor

Der Satz von Cantor wurde, wie der Name schon andeutet von Georg Cantor bewiesen.[Cantor:1962]. Cantor bewies mit einem Diagonalargument, dass die Menge aller Teilmengen der natürlichen Zahlen grösser ist als die Menge der natürlichen Zahlen. Dass eine Menge grösser ist als eine andere Menge, bedeutet dass es keine eineindeutige Abbildung zwischen den Mengen gibt. Diesen Beweis erweiterte er auf alle Mengen. Die Potenzmenge (Menge aller Teilmengen) einer Menge ist immer grösser als die Menge selbst. Wenn wir diesen Satz ernst nehmen, ihn also nicht sofort skolemisieren, dann sagt er dass es verschieden grosse Unendlichkeiten gibt. Insbesondere ist die Menge der reellen Zahlen grösser als die Menge der natürlichen bzw. die Menge der rationalen Zahlen.

Eines der wichtigsten Resultate der Mengentheorie führt es auch gleich zu einem Paradoxon, wenn wir uns fragen ob auch die Potenzmenge der Menge aller Mengen grösser ist als die Menge aller Mengen. Eine Lösung besteht darin, die Menge aller Mengen nicht als Menge zu betrachten und Mengen von echten Klassen zu unterscheiden.

Der erste, anschaulichere Beweis gilt nur für abzählbare Mengen, also insbesondere für die Menge der natürlichen Zahlen. Eine Menge heisst abzählbar, wenn sich die Elemente der Menge vollständig auflisten lassen. An diesem Beweis wird auch klar werden, warum diese Beweismethode den Namen "Diagonalisierung" bekommen hat.

Theorem 2.1.1 *(Satz von Cantor für abzählbar unendliche Mengen) Sei A eine abzählbar unendliche Menge. Es gibt keine bijektive Abbildung zwischen A und $\mathfrak{P}(A)$, der Potenzmenge von A.*

Beweis 2.1.1 *Angenommen wir haben eine Funktion h die eine abzählbar unendliche Menge A so auf ihre Potenzmenge $\mathfrak{P}(A)$ abbildet, dass jede Menge in $\mathfrak{P}(A)$ getroffen wird (Die Annahme der Surjektivität von h genügt für den Beweis).Angenommen die Funktion h sei als zweidimensionale Tabelle gegeben.*

x	$a \in h(x)$	$b \in h(x)$	$c \in h(x)$	$d \in h(x)$	$\ldots$
a	**0**	*1*	*1*	*0*	$\ldots$
b	*0*	**0**	*0*	*0*	$\ldots$
c	*1*	*1*	**0**	*0*	$\ldots$
d	*0*	*0*	*0*	**0**	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Diese Tabelle ist so zu interpretieren, dass in der ersten Spalte alle Elemente der Menge A aufgelistet sind, während in den Zeilen abzulesen ist ob die Menge auf die ein Element x abgebildet wird ein bestimmtes Element enthält 1 oder nicht enthält 0. a wird auf $\{b, c, \ldots\}$, b auf $\{\ldots\}$, c auf $\{a, b, \ldots\}$ abgebildet. Jetzt betrachten wir die Menge, die durch die Diagonale dargestellt wird, wenn wir jene Elemente in deren Spalte 0 steht in die Menge aufnehmen und jene in deren Spalte 1 steht nicht. Dann wird kein Element der Menge A auf diese Diagonalmenge abgebildet. Also gibt es keine surjektive und daher auch keine bijektive Abbidung zwischen einer (abzählbar unendlichen) Menge und ihrer Potenzmenge. $\square$

Der allgemeine Beweis verzichtet auf die Darstellung der Funktion mittels einer Tabelle. Gezeigt wird, dass es eine bestimmte Art von Funktion, nämlich eine surjektive zwischen einer Menge und ihrer Potenzmenge nicht geben kann.

Theorem 2.1.2 *(Satz von Cantor) Sei A eine beliebige Menge. Es gibt keine bijektive (d.h. auch keine surjektive) Abbildung zwischen A und $\mathfrak{P}(A)$, der Potenzmenge von A.*

Beweis 2.1.2 *Angenommen die Funktion h*

$$h : A \to \mathfrak{P}(A)$$

sei surjektiv, d.h. jedes Element aus $\mathfrak{P}(A)$ ist das Bild eines Elements aus A. Um zu zeigen, dass es eine solche Funktion nicht geben kann, müssen wir ein Element von $\mathfrak{P}(A)$ finden, dass kein Bild eines Elements von A sein kann. Sei C die Teilmenge von A, also ein Element von $\mathfrak{P}(A)$ für die gilt:

$$C = Die\ Menge\ aller\ x,\ sodass : x \notin h(x).$$

Angenommen es gäbe ein $c \in A$ sodass:

$$h(c) = C.$$

Wäre dann $c \in h(c)$, dann wäre $c \notin h(c)$. Wäre aber $c \notin h(c)$, dann wäre $c \in h(c)$. Damit haben wir einen Widerspruch hergeleitet, und können folgern dass es eine solche Funktion h nicht geben kann.□

2.1.2 Das Diagonallemma

Um das Diagonallemma zu beweisen, benötigen wir den Begriff der Gödelnumerierung [Goedel:1986] oder Gödelcodierung und den Begriff eines konsistenten formalen Systems.

Das Konzept der Gödelnumerierung werde ich hinreichend genau ausarbeiten um zu klären was es heissen soll, dass ein Satz von sich selbst irgendetwas behauptet. Da Sätze ja nichts behaupten, und da ich diese metaphorische Redeweise der Einfachheit halber später öfter verwenden werde, mir aber keinen Kategorienfehler nachsagen lassen will, werde ich diese Redeweise in diesem Abschnitt rechtfertigen.

Sei **S** ein formales System in der Sprache L, dass die Arithmetik enthält, Σ die Menge der Zeichenketten von L. Eine Gödelnumerierung ist eine Funktion $\ulcorner\urcorner$ von der Menge Σ in die natürlichen Zahlen N.

$$\ulcorner\urcorner : \Sigma \to N$$

Diese Funktion hat die Eigenschaft, dass ausgehend von einer natürlichen Zahl die Zeichenkette bestimmt werden kann, deren Bild die Zahl ist. Die Funktion realisiert also eine Codierung die auch decodiert werden kann. Sei $\sigma \in \Sigma$. $\ulcorner\sigma\urcorner$ wird Gödelnummer oder Gödelzahl von σ genannt

In **S** kann eine Gödelnumerierung für L definiert werden. Zuerst werden den Grundzeichen von L eineindeutig natürliche Zahlen zugeordnet.

$$''0''..1, \text{``}\neg''...5, \text{``}\vee''...7, \ldots$$

Dann werden die endlichen Folgen von natürlichen Zahlen eineindeutig auf die natürlichen Zahlen abgebildet, indem der Reihe $n_1, n_2, n_3 \ldots, n_k$ die natürliche Zahl

$$2^{n_1} * 3^{n_2} * 5^{n_3} * \ldots * p_k^{n_k}$$

zugeordnet wird, wobei p_k die k-te Primzahl bedeutet. Diese Zuordnung kann als Codierung verwendet werden, weil jede Zahl eindeutig in ihre Primzahlfaktoren n_i zerlegt werden kann.[1]

Wenn wir die Begriffe "Formel von L", "Beweis in **S**" usw. in **S** definieren können, dann können wir zahlentheoretische Aussagen als Aussagen über Formeln interpretieren. Um den Begriff "Formel von L" in **S** zu definieren, müssen wir ein zahlentheoretisches Prädikat finden, dass auf eine Zahl genau dann zutrifft, wenn diese Zahl die Gödelnummer einer Formel von L ist.

Die Interpretation der zahlentheoretischen Prädikate erfolgt zwar in der Metasprache, haben wir aber einmal ein zahlentheoretisches Prädikat identifiziert, dass nur auf die Gödelnummern von Formeln zutrifft, dann können wir sagen , dass **S** den Begriff "Formel von L" enthält.

Die Definition des Begriffes "Formel von L" in **S** werde ich jetzt etwas genauer skizzieren.[2]

Das Prädikat "Formel von L" in **S** zu definieren, bedeutet eine zahlentheoretische Formel ψ anzugeben, sodass für alle Zeichenketten σ in Σ gilt:

$$(\sigma \textit{ ist eine Formel }) \textit{ ist wahr } \leftrightarrow \mathbf{S} \vdash \psi(\ulcorner\sigma\urcorner)$$

[1] Siehe [Goedel:1986], S 156.

[2] Die Art der Codierung, und der Beweis des Diagonallemmas wurden von [Smorinski:1977] übernommen.

und

$$(\sigma \ ist\ eine\ Formel\ )\ ist\ falsch \leftrightarrow \mathbf{S} \vdash \neg\psi(\ulcorner\sigma\urcorner)$$

Eine zahlentheoretische Relation die auf diese Weise definiert werden kann, heisst "ziffernweise repräsentiert"("numeralwise represented") in **S**.[3]

Um das Prädikat "Formel von L" zu definieren, geben wir eine Funktion F an, sodass:

$$(\sigma\ ist\ eine\ Formel)\ ist\ wahr\ \leftrightarrow \mathbf{S} \vdash (F(\ulcorner\sigma\urcorner) = 1)$$

und

$$(\sigma\ ist\ eine\ Formel)\ ist\ falsch \leftrightarrow \mathbf{S} \vdash \neg(F(\ulcorner\sigma\urcorner) = 1)$$

F wird folgendermassen definiert:

$$F(x) = \begin{cases} 0 & \text{wenn } x \text{ eine Atomformel ist,} \\ 0 & \text{wenn } x \text{ eine negierte Formel ist,} \\ 0 & \text{wenn } x \text{ ein Konditional ist,} \\ 0 & \text{wenn } x \text{ eine allquantifizierte Formel ist,} \\ 1 & \text{anderenfalls.} \end{cases}$$

Das ist zwar kein Beweis, dass der Begriff "Formel von L" in **S** definiert werden kann, es sollte aber ausreichen um zu verstehen, wie eine solche Codierung durchgeführt werden kann. Kann man den Begriff "Beweis in **S**" in **S** definieren, dann sagen wir, dass **S** seine eigene Syntax codieren kann. In **S** können Aussagen über **S** gemacht werden.

Um das Diagonallemma zu beweisen, gehen wir davon aus, dass **S** seine eigene Syntax codieren kann, und konsistent ist. Ein inkonsistentes formales System kann seine eigene Syntax nicht definieren, denn man könnte in einem solchen System z.b.von jeder Zeichenkette beweisen dass sie eine Formel ist. Unter einem konsistenten formalen System **S** verstehen wir eine Menge von Sätzen der Sprache L die unter der Ableitbarkeitsrelation abgeschlossen ist. Wenn $\mathbf{S} \vdash \phi$, dann $\phi \in \mathbf{S}$. Konsistent bedeutet, dass nicht alle Sätze von L in **S** sind. Wir sagen, dass ein System **S** die Arithmetik enthält, wenn die Sprache von **S** die Sprache der Arithmetik enthält $(+, *, 0, 1)$, und **S** alle Axiome der Arithmetik enthält. Das Enthaltensein könnten wir auch anders interpretieren, etwa über eine Abbildung der Arithmetik in **S**.

[3] Alle sogenannten primitiv rekursiven Relationen und Funktionen können auf diese Weise in **S** repräsentiert werden, und weil für die Codierung der Syntax von **S** nur primitiv rekursive Relationen und Funktionen gebraucht werden, kann in **S** die Syntax von **S** definiert werden.

In einem formalen System, dass die Arithmetik enthält, können wir eine Gödelnumerierung und die gesamte Syntax dieses Systems definieren. Was ein Satz oder ein Term ist können wir im System selbst definieren. Wir können ein Prädikat $Satz(x)$ der Sprache L angeben, dass von einer natürlichen Zahl genaudann erfüllt wird, wenn sie Gödelnummer eines Satzes ist.

Zur Codierung der Syntax des Systems **S** gehört auch, dass wir eine zweistellige Funktion *sub* definieren die die Gödelnummer eines einstelligen Prädikats und einen geschlossenen Term als Argument erhält, und die Gödelnummer des Satzes der entsteht wenn das Zahlzeichen der Zahl auf die der Term referiert überall wo die Variable frei vorkommt eingesetzt wird, als Wert liefert. Was ich hier nicht beweisen kann ist, dass die Gödelnumerierung die Funktion *sub* usw. primitiv rekusive Funktionen sind und in einem formalen System das die Arithmetik enthält definiert werden können. Das diese Funktion in **S** definiert werden kann heisst, dass

$$\mathbf{S} \vdash sub(\ulcorner\phi(x)\urcorner, t) = \ulcorner\phi(n)\urcorner$$

und

$$\mathbf{S} \vdash t = n$$

ableitbar sind für $\phi(x)$ und t. t ist ein Term, der eine Zahl bezeichnet. Für diese Zahl gibt es einen Standardnamen n. Das Ergebnis der *sub*-Funktion liefert einen Satz in den die Gödelnummer des Standardnamens n eingesetzt wird.

Theorem 2.1.3 *(Diagonallemma). Sei* **S** *ein konsistentes formales Sytem, welches die Arithmetik enthält,* $\phi(x)$ *ein einstelliges Prädikat der Sprache von* **S**. *Dann gibt es einen Satz* ψ*, sodass:*

$$\mathbf{S} \vdash \psi \leftrightarrow \phi(\ulcorner\psi\urcorner).$$

Das Diagonallemma sagt, dass es für jedes Prädikat ϕ einen Satz gibt, der genau dann wahr ist, wenn seine Gödelnummer das Prädikat ϕ erfüllt. ϕ kann aber auch als ein Prädikat von Zeichenketten interpretiert werden. Dann behauptet ψ, eine Zeichenkette zu sein, die dieses Prädikat erfüllt.

Beweis 2.1.3 *Sei* $\phi(x)$ *ein beliebiges einstelliges Prädikat von* **S**. $\theta(x) \leftrightarrow \phi(sub(x, x))$ *nennen wir die Diagonalisierung von* ϕ. *Sei* m *die Gödelnummer der Diagonalisierung von* ϕ, $m = \ulcorner\theta(x)\urcorner$. *Sei* ψ *der Satz, den man erhält* *wenn man in die Diagonalisierung von ϕ die Gödelnummer der Diagonalisierung von ϕ einsetzt, $\psi = \theta(m)$. Dann ist*

$$\psi \leftrightarrow \theta(m) \leftrightarrow \phi(sub(m, m))$$

$$\leftrightarrow \phi(sub(\ulcorner\theta(x)\urcorner, m))$$

Die Funktion sub liefert als Wert die Gödelnummer des Satzes, den man erhält wenn man in $\theta(x)$ m einsetzt. Der Standardname der Zahl die m bezeichnet ist m. Darum liefert die Substitution

$$\leftrightarrow \phi(\ulcorner\theta(m)\urcorner) \leftrightarrow \phi(\ulcorner\psi\urcorner).$$

□

Setzt man für $\phi(x)$ $\neg Beweisbar_{\mathbf{S}}(x)$ ein , dann ist es nicht mehr schwierig Gödels erstes Unvollständigkeitstheorem zu beweisen. [Smorinski:1977]

In **S** können wir also zu jedem Prädikat eine Äquivalenz ableiten, die uns sagt das es einen Satz gibt der von sich selbst behauptet, dass er dieses Prädikat erfüllt. Das meine ich, wenn ich sage, dass wir alle selbstreferentiellen Sätze ausdrücken können. Weil die Sprache der Arithmetik eine sehr einfache Sprache ist und wir in ihr schon jeden selbstreferentiellen Satz ausdrücken können, kann eine Lösung der Lügnerparadoxie nicht darin bestehen die Selbstreferentialität einfach auszuschliessen. Auch jede formale Wahrheitstheorie muss die Existenz selbstreferentieller Sätze berücksichtigen.

Wir hätten auch intuitiv dafür argumentieren können, dass es selbstreferentielle Sätze gibt. "Dieser Satz hat fünf Wörter" wäre ein Beispiel. Dabei wäre es aber nicht sicher ob die Selbstreferentialität innerhalb des Systems entsteht oder von ausserhalb kommt.

2.1.3 Die "Diagonalheit" des Diagonallemmas

Warum heisst dieses Lemma "Diagonallemma"? Welchen Zusammenhang gibt es zwischen der Verwendung der Methode der Diagonalisierung beim Satz von Cantor und dem Diagonallemma? Ich werde versuchen dass durch einen anschaulichen Beweis klarzumachen. Für diesen Beweis verwende ich dieselbe Methode die ich zum Beweis des Satzes von Cantor verwendet habe. Das Konzept der Gödelnumerierung wird ebenfalls verwendet.

Sei folgende Tabelle gegeben.

	$\ulcorner\phi_1(x)\urcorner$	$\ulcorner\phi_2(x)\urcorner$	$\ulcorner\phi_3(x)\urcorner$	$\ulcorner\phi_4(x)\urcorner$	$\dots$
$\phi_1(x)$	$\phi_1(\ulcorner\phi_1(x)\urcorner)$	$\phi_1(\ulcorner\phi_2(x)\urcorner)$	$\phi_1(\ulcorner\phi_3(x)\urcorner)$	$\phi_1(\ulcorner\phi_4(x)\urcorner)$	$\dots$
$\phi_2(x)$	$\phi_2(\ulcorner\phi_1(x)\urcorner)$	$\phi_2(\ulcorner\phi_2(x)\urcorner)$	$\phi_2(\ulcorner\phi_3(x)\urcorner)$	$\phi_2(\ulcorner\phi_4(x)\urcorner)$	$\dots$
$\phi_3(x)$	$\phi_3(\ulcorner\phi_1(x)\urcorner)$	$\phi_3(\ulcorner\phi_2(x)\urcorner)$	$\phi_3(\ulcorner\phi_3(x)\urcorner)$	$\phi_3(\ulcorner\phi_4(x)\urcorner)$	$\dots$
$\phi_4(x)r$	$\phi_4(\ulcorner\phi_1(x)\urcorner)$	$\phi_4(\ulcorner\phi_2(x)\urcorner)$	$\phi_4(\ulcorner\phi_3(x)\urcorner)$	$\phi_4(\ulcorner\phi_4(x)\urcorner)$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

In der ersten Spalte sind alle Prädikate einer Sprache L aufgelistet. In der ersten Zeile alle Gödelnummern der Prädikate. Im Inneren der Tabelle stehen die Sätze die man erhält wenn man in die Prädikate die entsprechenden Gödelnummern einsetzt. In der Diagonale stehen die Sätze die man erhält wenn man in die Prädikate ihre eigenen Gödelnummern einsetzt.

Angenommen wir suchen einen Satz der von sich selbst behauptet das Prädikat $\phi_1(x)$ zu erfüllen. Nach dem Diagonallemma gibt es so einen Satz für jedes Prädikat $\phi_1(x)$.

$$\phi_1(\ulcorner\phi_1(x)\urcorner)$$

kann nicht der gesuchte Satz sein, denn dieser Satz behauptet lediglich dass $\ulcorner\phi_1(x)\urcorner$ das Prädikat $\phi_1(x)$ erfüllt..

Also suchen wir entlang der ersten Spalte weiter, bis wir das Prädikat $\phi_1(sub(x,x)$ finden. Wenn wir $sub(x,x)$ in L formalisieren können, dann finden wir dieses Prädikat auch auf der Liste. An dieser Stelle finden wir auf der Diagonale den Satz

$$(*)\phi_1(sub(\ulcorner\phi_1(sub(x,x)\urcorner,\ulcorner\phi_1(sub(x,x)\urcorner)$$

.Sei $m=\ulcorner\phi_1(sub(x,x)\urcorner$. Dann können wir (*) auch schreiben als:

$$\phi_1(sub(\ulcorner\phi_1(sub(x,x)\urcorner,m).$$

Wenn wir jetzt die Substitution durchführen (d.h. den Wert der Funktion sub mit den Argumenten $\ulcorner\phi_1(sub(x,x)\urcorner$ und m errechnen), dann erhalten wir eine Zahl n, sodass

$$n=\ulcorner\phi_1(sub(m,m)\urcorner$$

und wir können (*) schreiben als

$$\phi_1(n)$$

Da

$$\phi_1(sub(\ulcorner\phi_1(sub(x,x)\urcorner,m)\leftrightarrow\phi_1(sub(m,m)\leftrightarrow\phi_1(n)$$

ist, haben wir den gesuchten Satz gefunden.

$\phi_1(x)$ ist zwar nur ein Prädikat der Zahlentheorie, und $\phi_1(n)$ ist ein zahlentheoretischer Satz. Da wir aber die Begriffe Formel, Satz und Beweis in der Zahlentheorie definiert haben, und alle Formeln in die natürlichen Zahlen abgebildet haben, können wir jede Aussage über Zahlen als eine Aussage über Formeln interpretieren.

$\phi_1(n)$ "behauptet" dann, dass der Satz mit der Gödelnummer n das Prädikat ϕ_1 erfüllt. Aber n ist gerade die Gödelnummer von $\phi_1(n)$. Da sich diese Konstruktion für jedes beliebige Prädikat durchführen lässt, haben wir das Diagonallemma ohne Beschränkung der Allgemeinheit bewiesen. Um die Konstruktion durchzuführen haben wir nur das Konzept der Gödelnumerierung (sowie die Funktion *sub*) verwendet.

Der Zusammenhang mit der Methode der Diagonalisierung dürfte auch klar geworden sein.

2.2 Undefinierbarkeit der Wahrheit in einem formalen System

Wenn wir von der Undefinierbarkeit der Wahrheit in einem formalen System **S** sprechen, dannn meinen wir, dass die Menge der wahren Sätze von **S** nicht in **S** definiert werden kann. Es gibt kein einstelliges Prädikat von **S** dass genau auf die wahren Sätze von **S** zutrifft.

Alfred Tarski [Tarski:1939] hat als erster bewiesen, dass eine Definition des Wahrheitsprädikats in einem formalen System dass die Arithmetik enthält[4], nicht gegeben werden kann. Bevor ich mich etwas genauer mit Tarski's Beweis beschäftige, werde ich einen einfacheren Beweis angeben, bei dem das Diagonallemma verwendet wird.

2.2.1 Beweis mittels Diagonallemma

Nehmen wir wieder an **S** sei ein fomales System, dass die Arithmetik enthält. Weiters sei $Wahr(x)$ ein Prädikat, sodass für jeden Satz ϕ gilt:

$$(T - Schema)\ \mathbf{S} \vdash Wahr(\ulcorner \phi \urcorner) \leftrightarrow \phi$$

Das Diagonallemma sagt uns, dass es einen Satz ψ gibt der von sich selbst behauptet, dass er nicht wahr ist. Dieser Satz ist die formalisierte Variante des Lügnerparadoxons.

$$\mathbf{S} \vdash \neg Wahr(\ulcorner \psi \urcorner) \leftrightarrow \psi$$

Aus dem T-Schema folgt:

$$\mathbf{S} \vdash Wahr(\ulcorner \psi \urcorner) \leftrightarrow Wahr(\ulcorner \psi \urcorner)$$

Mittels eines einfachen aussagenlogischen Schlusses erhalten wir:

$$\mathbf{S} \vdash \neg Wahr(\ulcorner \psi \urcorner) \leftrightarrow Wahr(\ulcorner \psi \urcorner)$$

Damit haben wir aus unserer Ausgangsprämisse einen Widerspruch abgeleitet. Da wir in der klassischen Logik arbeiten, können wir mittels einer *reductio ad absurdum* folgenden Satz herleiten :

$$\mathbf{S} \nvdash Wahr(\ulcorner \phi \urcorner) \leftrightarrow \phi$$

[4]Es genügt, dass bestimmte Codierungen in **S** durchgeführt werden können. Tarski merkt an:

> *We assume it to be understood that metalogical statements about the system* **S** *can, at least in part, be formalized, or rather interpreted, in the system* **S** *itself.* [Tarski:1939], S 106

> *This was found by K.Gödel and the present author independently of one another.* [Tarski:1939], S 106, Anmerkung 5

> *...zwischen den Ausdrücken (des formalen Systems* **S***) und den natürlichen Zahlen lässt sich eine eineindeutige Zuordnung durchführen, wobei man den Operationen an Ausdrücken Operationen an Zahlen von denselben formalen Eigenschaften zuordnen kann.* [Tarski:1971], S 478

In einem formalen System, in dem das Diagonallemma gilt, kann das T-Schema also nicht bewiesen werden. Wenn wir das T-Schema als Bedingung akzeptieren, die ein Wahrheitsprädikat erfüllen muss, dann gilt das obige Resultat für jedes Wahrheitsprädikat.

Was bedeutet dieses Resultat für die Semantik von natürlichen Sprachen? Ist es überhaupt direkt anwendbar?

Da wir bei dem Beweis das Konzept der Gödelnumerierung verwendet haben, können wir das Resultat nicht direkt auf natürliche Sprachen anwenden. Dazu müsste zuerst gezeigt werden, dass solche Numerierungen auch in der natürlichen Sprache vorkommen. Der Hinweis, dass die natürliche Sprache die Arithmetik enthält, und damit auch Gödelnumerierungen ist sicher zuwenig, denn wenn wir in der Sprache über die Sprache reden, dann verwenden wir keine Codierung via die natürlichen Zahlen.

Das Diagonallemma sagt uns zwar nichts darüber wie Selbstreferenz in der natürlichen Sprache funktioniert, es sagt aber einiges über die Möglichkeiten der natürlichen Sprache selbstreferentielle Sätze zu bilden. Es ist möglich jeden selbstreferentiellen Satz zu bilden.

Dass das Wahrheitsprädikat auch in der natürlichen Sprache undefinierbar ist, zeigt sich wenn man obigen Beweis in natürliche Sprache übersetzt. Eine Übersetztung die möglich ist, weil "der Lügner" in der natürlichen Sprache formuliert werden kann.

2.2.2 Tarski's Beweis

Obwohl der Originalbeweis von Tarski [Tarski:1939] von der Grundidee her derselbe ist, unterscheidet er sich doch in einigen Punkten vom vorigen Beweis. Diese Punkte möchte ich jetzt herausarbeiten.

Auch Tarski's Beweis gilt für ein beliebiges formales System **S** dass die Arithmetik enthält.[5] Seien $x, y, z, \ldots$ Variablen für Formeln, Se die Menge der Sätze, D die Menge der beweisbaren Formeln, und $\tilde{z}$ die Negation von z. Die Undefinierbarkeit der Wahrheit beweist Tarski folgendermassen. Zuerst beweist er folgendes Theorem.[6]

---

[5] Tarski's Bedingung an das formale System:

> *that the laws of the system suffice for the construction of the arithmetic of natural numbers.* [Tarski:1939], S 105

[6] Im Artikel [Tarski:1939], S 107 erwähnt Tarski nur, dass er dieses Lemma schon wo-

Theorem 2.2.1 *There is a statement $z \in Se$ such that $\tilde{z} \leftrightarrow z_{(E)} \in D$.*

In Analogie zu unserer obigen Schreibweise des Diagonallemmas könnten wir Theorem 2.2.1 folgendermassen übersetzen.

$$(*)\mathbf{S} \vdash \neg\phi \leftrightarrow \psi(\ulcorner\phi\urcorner)$$

Zu jedem einstelligem Prädikat gibt es also einen Satz, der von sich behauptet dieses Prädikat nicht zu erfüllen. Dass dieses Theorem aus dem Diagonallemma herleitbar ist, kann man leicht sehen. Da das Diagonallemma für alle Prädikate gilt, gilt es auch für alle negierten Prädikate. Es gilt also:

$$\mathbf{S} \vdash \phi \leftrightarrow \neg\psi(\ulcorner\phi\urcorner)$$

Durch einfache aussagenlogische Umformung erhält man (*).

Ausserdem beweist Tarski folgendes Theorem:

Theorem 2.2.2 *IF $E \subset Tr$, and $\tilde{z} \leftrightarrow z_{(E)} \in D$, then $z \in Tr$, $z \notin E$ and $\tilde{z} \notin E$.*

Wenn E eine Teilmenge der wahren Sätze ist, und es einen Satz gibt, der von sich selbst behauptet nicht in E enthalten zu sein, dann ist dieser Satz wahr, und weder er noch seine Negation sind in E. Übersetzt in unsere Notation:

Theorem 2.2.3 *Sei E ein Prädikat, dass nur auf wahre Formeln, aber nicht notwendigerweise auf alle wahren Formeln zutrifft. Wenn $\mathbf{S} \vdash \neg\phi \leftrightarrow E(\ulcorner\phi\urcorner)$, dann ist $Wahr(\ulcorner\phi\urcorner)$ (kann nicht in S bewiesen werden), $\neg E(\ulcorner\phi\urcorner)$ und $\neg E(\ulcorner\neg\phi\urcorner)$.*

Beweis 2.2.1 *Angenommen $\neg Wahr(\ulcorner\phi\urcorner)$. Dann folgt aus der Vollständigkeit des Wahrheitsprädikats, dass $Wahr(\ulcorner\neg\phi\urcorner)$. Da alle beweisbaren Sätze von S auch wahr sind, folgt dass $Wahr(\neg\phi \rightarrow E(\ulcorner\phi\urcorner))$. Da "Wahr" bezüglich logischer Folgerung abgeschlossen ist, können wir aus den letzten beiden Tatsachen $Wahr(\ulcorner E(\ulcorner\phi\urcorner)\urcorner)$ herleiten. Mittels T-Schema und der Tatsache, dass "E" ein Teilprädikat von "Wahr" ist, können wir $Wahr(\ulcorner\phi\urcorner)$ herleiten. Das widerspricht aber unserer Annahme.*

Es folgt auch sofort, dass $\neg E(\ulcorner\neg\phi\urcorner)$. Wäre $E(\ulcorner\neg\phi\urcorner)$, dann wäre $Wahr(\ulcorner\neg\phi\urcorner)$, und dann wäre "Wahr" inkonsistent.

Aus unseren Annahmen können wir $\phi \rightarrow \neg E(\ulcorner\phi\urcorner)$ ableiten. Wenn $Wahr(\ulcorner\phi\urcorner)$ dann folgt $Wahr(\ulcorner\neg E(\ulcorner\phi\urcorner)\urcorner)$. Da "Wahr" konsistent ist, folgt daraus $\neg Wahr(\ulcorner E(\ulcorner\phi\urcorner)\urcorner)$, und $\neg E(\ulcorner\phi\urcorner)$.

ϕ hat also alle erforderlichen Eigenschaften. □

anders bewiesen hat, und zwar in [Tarski:1971].

Aus den Theoremen 2.2.1 und 2.2.2 folgt:

Theorem 2.2.4 *IF $E \subset Tr$, then there is a statement which is undecidable in E, i.e., a $z \in Se$ such that $z \notin E$ and $\tilde{z} \notin E$.*

Zum Theorem 2.2.4 merkt Tarski an:

> *..it follows at once that it is impossible to construct such a definition of Tr (=die Menge aller wahren Sätze) as could be formalized in S. If the contrary where the case undecidable statements could be found in Tr, and this would be incompatible with 2.2.5*

Theorem 2.2.5 besagt, dass die Menge Tr vollständig ist. D.h. wenn ein Satz nicht in Tr enthalten ist, dann ist seine Negation in Tr enthalten.

Theorem 2.2.5 *If $x \in Se$ and $x \notin Tr$, then $\tilde{x} \in Tr$.*

Jede Teilmenge von Tr, die im System S konstruiert werden kann, ist also notwendigerweise unvollständig. Könnte die Menge Tr in S konstruiert werden, würde das der Vollständigkeit des Prädikats Tr widersprechen. Die Annahme, man könnte Tr in S definieren widerspricht dem Theorem 2.2.5. Wenn wir Theorem 2.2.5 wieder in unsere Notation übersetzen, erhalten wir:

$$\neg Wahr(\ulcorner\phi\urcorner) \rightarrow Wahr(\ulcorner\neg\phi\urcorner)$$

Mittels *modus tollendo tollens* erhalten wir:

$$(**)\neg Wahr(\ulcorner\phi\urcorner) \leftrightarrow Wahr(\ulcorner\neg\phi\urcorner)$$

(**) können wir einfach aus dem T-Schema herleiten. Im wesentlichen unterscheidet sich Tarski's Beweis also nicht vom Beweis mittels Diagonallemma. Mit dem Theorem 2.2.4 beweist er nur explizit, dass es zu jedem möglichen Wahrheitsprädikat in S unentscheidbare Sätze gibt.

2.2.3 Frege's Argument

Gottlob Frege [Frege:1966] bringt ein anderes Argument für die Undefinierbarkeit der Wahrheit. Er schreibt:

> *" So scheitert aber auch jeder andere Versuch, das Wahrsein zu definieren. Denn in einer Definition gäbe man gewisse Merkmale an. Und bei der Anwendung auf einen besonderen Fall käme es dann immer darauf an. ob es wahr wäre, daß diese Merkmale zuträfen."* [7]

Eine Wahrheitsdefinition wäre nach Frege also notwendigerweise zirkulär. Und das stimmt ja auch. Wenn wir versuchen die Wahrheit eines Satzes zu definieren, indem wir andere Eigenschaften des Satzes verwenden, dann kann immer die Frage gestellt werden, ob es wahr ist dass der Satz diese anderen Eigenschaften hat, bzw. hängt die Wahrheit des Satzes davon ab, ob er diese anderen Eigenschaften hat. Angenommen wir versuchen das Wahrheitsprädikat folgendermassen zu definieren.

$$Wahr(\ulcorner\phi\urcorner) \leftrightarrow F(\ulcorner\phi\urcorner))$$

Wenn das Prädikat F aber auf ϕ zutrifft, dann muss es wahr sein, dass es so ist. Und dann haben wir das Wahrheitsprädikat, dass wir definieren wollten im Definiens verwendet.

$$Wahr(\ulcorner\phi\urcorner) \leftrightarrow Wahr(\ulcorner F(\ulcorner\phi\urcorner)\urcorner)$$

Das T-Schema von Tarski umgeht diese Schwierigkeit dadurch, dass gar keine Eigenschaften des Satzes angegegen werden die notwendige und hinreichende Bedingungen dafür sind dass der Satz wahr ist.

Darum ist es auch so schwierig die Bedeutung des T-Schemas in natürlicher Sprache anzugeben. Wenn ich einen bestimmten Satz habe z.b." Schnee ist weiss", dann besagt das T-Schema dass "Schnee ist weiss" wahr ist genau dann wenn Schnee weiss ist.

Was bedeutet das Schema aber allgemein, für alle Sätze? Versuche ich die Bedeutung des Schemas durch "Ein Satz ist wahr genau dann, wenn er behauptet werden darf" anzugeben, dann gebe ich eine Eigenschaft des Satzes , nämlich die "behauptet werden zu dürfen" im Definiens an.

2.3 Partielle Prädikate

Unsere bisherige Untersuchung bezüglich der Definierbarkeit bzw. Undefinierbarkeit des Wahrheitsprädikats in einem formalen System, könnte uns

[7] In Frege [Frege:1966], S.32

zu der Auffassung bringen, dass etwas mit unserem Verständnis des Wahrheitsprädikats nicht in Ordnung ist. Wir haben immer eine Definition eines totalen Prädikats gesucht. Entweder sollte das Wahrheitsprädikat auf einen Satz zutreffen, oder es sollte das Falschheitsprädikat auf einen Satz zutreffen, dass in diesem Fall die Negation des Wahrheitsprädikats wäre. Realistisch gesprochen, sollte jeder Satz entweder die Eigenschaft haben wahr zu sein, oder die Eigenschaft haben falsch zu sein. (Das "entweder oder" ist im exklusiven Sinn zu verstehen, d.h. die Disjunktion ist wahr genau dann wenn eines der beiden Disjunkte wahr ist) Wenn wir diese Vorraussetzung aufgeben, und versuchen würden ein partielles Wahrheitsprädikat zu definieren, könnten wir vielleicht mehr erreichen.

Die Logik eines partiellen Wahrheitsprädikats, lässt sich durch folgende Regel charakterisieren. Wenn wir festgestellt haben dass ein Satz nicht wahr ist, können wir nicht darauf schliessen dass er falsch ist.

In diesem Abschnitt möchte ich die Verwendung von partiellen Prädikaten und den Übergang von der klassischen zur mehrwertigen Logik motivieren. Einen Grund für die Einführung partieller Prädikate haben wir schon kennengelernt. Die Undefinierbarkeit eines totalen Wahrheitsprädikats.

Zuerst möchte ich einiges darüber sagen was es bedeuten kann, dass ein Prädikat nicht definiert ist. In der mehrwertigen Logik werden zusätzliche Wahrheitswerte eingeführt. Wenn ein Prädikat für einen bestimmten Gegenstand nicht definiert ist, dann hat der Satz den man erhält, wenn man in das Prädikat einen Namen des Gegenstandes einsetzt einen Wahrheitswert, der von 'Wahr' und 'Falsch' unterschieden wird. Nennen wir diesen Wahrheitswert 'undefiniert'. Auf diese Weise können beliebig viele Wahrheitswerte, mit unterschiedlichen Bedeutungen eingeführt werden.

Anschliessend möchte ich die Einführung der mehrwertigen Logik von der Seite der Rekursionstheorie her motivieren.

2.3.1 Typen von Nullwerten

C.J. Date gibt sieben verschiedene Typen von Nullwerten an [Date:1989]. Nullwerte erfüllen in der Datenbanktheorie und Praxis dieselbe Rolle, wie zusätzliche Wahrheitswerte in der mehrwertigen Logik.

Eine Datenbank besteht aus Tabellen. Eine Tabelle ist eine Matrix, deren Spalten als Attribute von Objekten interpretiert werden, und deren Zeilen als Objekte interpretiert werden. Die folgende Tabelle

Farbe	**Art**	Geruch	Geschmack
Gelb	**Zitrone**		Sauer
Rot	**Apfel**	Blumig	Süss

Speichert Informationen über Früchte. Jede Zeile (jedes Objekt) besitzt einen Schlüssel, in diesem Fall das Attribut 'Art' durch den sie (es) innerhalb der Tabelle eindeutig identifiziert werden kann. Das Attribut 'Geruch' des Objekts 'Zitrone' hat einen Nullwert.

Übersetzt in eine mehrwertige Sprache bedeutet es, dass der Satz den man erhält, wenn man in das Prädikat 'Geruch' den Eigennamen 'Zitrone' einsetzt, den Wahrheitswert 'undefiniert' hat bzw. bezüglich seines Wahrheitswerts undefiniert ist.

Ein Nullwert ist eigentlich kein Wert, sondern bedeutet dass an dieser Stelle ein Wert fehlt. C.J. Date schreibt:

> *" "Joe's salary is null" is intended to mean that there is a position in the database for recording Joe's salary, but no value is recorded at that postion at the present time. Note immediately, therefore, that "null" is not a value; rather, it is a representation of the fact that there is no value.*[8]

Franz von Kutschera behauptet dasselbe für die "Wahrheitswerte" der mehrwertigen Logik.

> *Indeterminiertheit ist auch kein dritter Wahrheitswert, sondern eine Unbestimmtheit des Wahrheitswerts..*[9]

Die Unbestimmtheit des Wahrheitswerts bzw. die Undefiniertheit eines Satzes kann aber sehr verschiedenes bedeuten. C.J.Date hat in seinen Aufsätzen zur Datenbanktheorie die Vieldeutigkeit der "Nullwerte" herausgearbeitet. Date gibt folgende Bedeutungen von Nullwerten an.

1. Ein Attribut ist auf ein Objekt nicht anwendbar. Date bringt das Beispiel eines Angestellten, auf den das Attribut "Provision" nicht anwendbar ist, weil nur Angestellte bestimmter Abteilungen Provisionen erhalten. Da Date diese Art von Nullwert von Nullwerten unterscheidet, bei denen das Attribut einfach nicht definiert ist, könnte man diese Art von Nullwerten als Ergebnisse von Kategorienfehlern verstehen.

[8] [Date:1989], S 218.
[9] [Kutschera:1985], S 146.

Mit dem Begriff "Provision" ist der Begriff der Kategorie aber etwas zu tief angesetzt. Zu verstehen was eine Provision ist, ist sicher keine Bedingung der Möglichkeit von Verstehen überhaupt.

Wie zu entscheiden ist, ob ein Nullwert das Ergebnis eines Kategorienfehlers ist oder das Ergebinis einer Definition kann ich nicht sagen, und auch Date gibt Beispiele zur Unterscheidung an, die man beiden Typen von Wahrheitswerten zuordnen könnte.

Würden wir einen solchen Nullwert in eine mehrwertige Logik übersetzen, dann wäre die Logik der Junktoren durch folgende Regeln gegeben.

- Ist eines der Attribute A oder B wahr und eines NULL (nicht anwendbar) dann ist das zusammengesetzte Attribut A *und* B wahr.

  Das Attribut, dass nicht anwendbar ist, kann das zusammengestzte Attribut auf keinen Fall falsch machen, darum ist A *und* B wahr. (Ein Attribut ist wahr, bezüglich eines Gegenstands, wenn es auf diesen Gegenstand zutrifft.)

- Ist das Attribut A NULL, dann ist das Attribut *nicht* A auch NULL.

2. Der Wert ist nicht bekannt.Z.b. der Wert einer Funktion, der aufwendig zu berechnen ist, oder der gar nicht berechnet werden kann. Aus der Tatsache dass etwas nicht bekannt ist, lässt sich nicht folgern dass es möglich sein muss damit bekannt zu werden. Diese Folgerung beruht auf einem erkenntnistheoretischen Postulat, das nicht in der Logik der Nullwerte enthalten ist. Diese Art von Nullwerten ist in den meisten Abfragesprachen von Datenbanksystemen implementiert.

   Die Logik der Junktoren wird durch folgende Regeln bestimmt.

   - Wenn das Attribut A NULL (nicht bekannt) ist, und das Attribut B NULL ist, dann ist das Attribut A *und* B auch NULL. Ist eines von beiden falsch, dann ist das zusammengesetzte Attribut falsch. Das zusammengesetzte Attribut ist nur dann wahr, wenn beide Attribute wahr sind.

     Ein Attribut dass nicht bekannt ist kann falsch werden, darum kann auch das Attribut A *und* B falsch werden.

- Wenn das Attribut A NULL ist, dann ist das Attribut *nicht* A auch NULL.

3. Der Wert existiert nicht. Dieser Typ von Nullwert muss vom ersten Typ unterschieden werden. Für diese Unterscheidung benutzt C.J. Date ein Beispiel aus der Versicherungsbranche. Obwohl das Attribut "Sozialversicherungsnummer" auf jeden anwendbar ist, kommt es vor dass einige Menschen keine Sozialversicherungsnummer haben. Die beiden ersten Typen haben wir mittels Regeln für die aussagenlogischen Junktoren erklärt. Bei diesem Typ müssen wir die Prädikatenlogik verwenden.

   - Wenn das Attribut A NULL (es existiert kein Wert) ist, dann ist *es gibt kein* x *sodass* A wahr.

   Bei einem Kategorienfehler wäre diese Existenzaussage nicht wahr, sondern falsch.

4. Der Wert ist nicht definiert. Die Division durch Null ist ein Beispiel dafür. Ob ein Kategorienfehler vorliegt, oder ob ein Wert nicht definiert ist, ist meiner Meinung nach in diesem Kontext nicht zu entscheiden. Logisch ist eine solche Unterscheidung nicht unbedingt nötig, da für beide Arten von Wahrheitswerten dieselben Junktorenreglen verwendet werden können. Vergleiche Typ 1 und Typ 4.

5. Der Wert ist ungültig. Würde in einer Datenbank eine Null von diesem Typ eingetragen, dann würden Prozeduren gestartet werden die versuchen einen gültigen Wert zu finden. Eine Null dieses Typs ist eine "Übergangslösung". Ein Wert, der eingetragen wird bis ein besserer gefunden wird.

6. Der Wert wird nicht unterstützt. Aus operationalen Gründen kann es brauchbar sein einen solchen Nulltyp zu implementieren.

7. Der Wert ist die leere Menge. Anstatt dieses Werts kann man auch eine "0" verwenden.

8. Aus diesen Werten können jetzt unendlich viele Nullwerte generiert werden. Angenommen, wir haben n verschiedene Nullwerte. Dann können

wir einen neuen Nullwert einführen mit der Bedeutung "Es ist nicht bekannt ob das Attribut den Nullwert 1, den Nullwert 2, .. oder den Nullwert n hat". Mit diesem Nullwert kann diesselbe Konstruktion noch einmal durchgeführt werden.

Die in den folgenden Kapiteln verwendete mehrwertige Logik entspricht der zweiten Variante.

2.3.2 Partiell rekursive Funktionen

In diesem Abschnitt möchte ich den Zusammenhang zwischen partiellen Relationen und partiell rekursiven Funktionen erklären. Dieser Zusammenhang ist interessant, weil der Begriff der partiell rekursiven Funktion ein spezieller formaler Begriff ist.

Es wurde bewiesen, dass viele andere Formalismen mittels derer Funktionen definiert werden können dieselbe Klasse von Funktionen (partiell rekursive Funktionen) definieren. Das führte zu der Vermutung, dass alle Funktionen die durch eine effektive Prozedur berechnet werden können partiell rekursiv seien.[10] Da der Begriff "ist durch eine effektive Prozedur berechenbar" kein formaler Begriff ist, kann diese Vermutung nicht durch einen formalen Beweis bestätigt werden. Obwohl diese Äquivalenz nicht bewiesen werden kann, gelten partiell rekursive Funktionen als Standard dafür was definiert werden kann.

Der formale Begriff der partiell rekursiven Funktion muss daher um den Begriff der partiell rekursiven Relation erweitert werden. In der klassische Modelltheorie wird, wie wir gesehen haben eine Relation als eine Menge von Tupeln verstanden. Die Elemente die in der Menge sind stehen in der Relation zueinander, die anderen nicht . Eine Relation kann aber auch als eine Funktion von einem Wertebereich in den Bereich der Wahrheitswerte $\{\mathbf{t}, \mathbf{f}\}$ verstanden werden.

Wird ein Tupel aus dem Wertebereich auf $\mathbf{t}$ abgebildet, dann stehen die Elemente des Tupels in der entsprechenden Relation zueinander, wird es auf $\mathbf{f}$ abgebildet dann stehen sie nicht in der Relation zueinander. Um auch partielle Relationen definieren zu können, müssen wir über die klassische Modelltheorie hinausgehen. Wird ein Tupel auf keinen Wert abgebildet, d.h. die

[10] Die sogenannte Church'sche These. Für eine Analyse der verschiedenen Thesen, die unter diesem Namen diskutiert werden siehe John T. Kearns. [Kearns:1997]

Funktion ist für dieses Element undefiniert, dann ist nichts über die Beziehung der Elemente bekannt.

Jede partiell rekursive Funktion P,

$$P : D^n \to \{\mathbf{t}, \mathbf{f}\}$$

kann als Repräsentation einer partiellen Relation interpretiert werden. P repräsentiert dieselbe Relation wie die Menge A, wenn für alle $a \in D^n$, für die $P(a)$ definiert ist gilt, dass

$$a \in A \leftrightarrow P(a) = \mathbf{t}.$$

Das führt zu folgender Definition.

Definition 2.3.1 *Eine Relation ist* **partiell rekursiv**, *wenn sie durch eine* **partiell rekursive** *Funktion repräsentiert werden kann.*

Wir können somit partielle Prädikate mittels partiell rekursiver Funktionen definieren, ohne eine dritten Wahrheitswert zu verwenden. Es gibt auch ein prinzipielles Problem, wenn ein zusätzlicher Wahrheitswert verwendet wird. Es können dann zwar partielle Relationen mittes totaler Funktionen repräsentiert werden, es fallen aber partielle Funktionen aus dieser Definition heraus, weil nicht für jede partielle Funktion und jedes Argument entschieden werden kann ob die Funktion für dieses Argument definiert ist.

Kapitel 3

Definierbare Wahrheitsprädikate

3.1 Wahrheitsprädikate für Teilsprachen

Wie wir gesehen haben sind die Möglichkeiten ein Wahrheitsprädikat für eine Sprache in dieser Sprache selbst zu definieren, begrenzt. Ein Wahrheitsprädikats für ein formales System **S** in diesem formalen System zu definieren, bedeutet, dass für jeden Satz ϕ in der Sprache von **S** gilt:

$$\mathbf{S} \vdash Wahr(\ulcorner \phi \urcorner) \leftrightarrow \phi$$

"Wahr" ist nur die Abkürzung für einen Prädikatausdruck der Sprache von **S**. Der Einfachheit halber werde ich mich im folgenden Kapitel nur mit dem formalen System der Arithmetik (**PA** für Peano-Arithmetik) beschäftigen.

Wenn wir unsere Ansprüche an ein Wahrheitsprädikat zurückschrauben, und wir dass T-Schema nicht für alle Sätze der Sprache von **PA** beweisen wollen, werden wir dann erfolgreich sein. Wir werden.

Für bestimmte Teilmengen von Sätzen von **PA** können wir ein Wahrheitsprädikat definieren.

3.1.1 Wahrheitsprädikate für Satzmengen mit beschränkter logischer Komplexität

Die Sätze von **PA** können wir nach ihrer logischen Komplexität ordnen. Die logische Komlexität einer Formel ϕ, $kompl(\ulcorner \phi \urcorner)$ wird folgendermassen definiert.[1]

Definition 3.1.1 *Wenn ϕ eine Atomformel ist, dann ist*

$$kompl(\ulcorner\phi\urcorner) = 0.$$

Ist $kompl(\ulcorner\phi\urcorner) = n,\ n \in N$ dann ist

$$kompl(\ulcorner\forall x\phi\urcorner) = kompl(\ulcorner\neg\phi\urcorner) = n + 1.$$

Ist $max(kompl(\ulcorner\phi\urcorner), kompl(\ulcorner\psi\urcorner)) = n,\ n \in N$ dann ist

$$kompl(\ulcorner\phi \vee \psi\urcorner) = n + 1.$$

Sei L die Sprache von **PA**, $clterm(x)$ ein Prädikat, dass auf eine Zahl zutrifft, wenn sie Gödelnummer eines geschlossenen Terms ist, $val(x)$ eine Funktion die als Argument eine Gödelnummer eines geschlossenen Terms erhält, und als Wert die Zahl liefert, die der Term denotiert, $*_1$ ein Verknüpfungsoperator, der als Argumente Gödelnummern von Ausdrücken $a, b, c, ..n$ erhält und als Wert die Gödelnummer des zusammengesetzten Ausdrucks $abc..n$ liefert und $Satz(x)$ ein entsprechendes zahlentheoretisches Prädikat für die Sätze von L. "$=_{def}$" bedeutet, dass der auf der linken Seite des Zeichens stehende Ausdruck nur eine Abkürzung für den auf der rechten Seite stehen Ausdruck ist. $clterm(x)$, $val(x)$, $*_1$ und $Satz(x)$ sind zahlentheoretische Funktionen bzw. Prädikate die in **PA** definierbar sind. Ausserdem verwenden wir eine Variation der Funktion $sub(x, y)$, nämlich $sub_1(x, y, z)$, die als Wert die Gödelnummer des Ausdrucks liefert, den man erhält wenn man im Ausdruck mit der Gödelnummer x alle freien Vorkommnisse der Variablen mit der Gödelnummer y durch das Zahlzeichen für z ersetzt.

Definition 3.1.2 *Das Wahrheitsprädikat für die Sätze mit logischer Komplexität 0, $Wahr_0(\ulcorner\phi\urcorner)$ wird folgendermassen definiert:*

$$Wahr_0(\ulcorner\phi\urcorner) =_{def} \exists t \exists s (clterm(t) \wedge clterm(s) \wedge \ulcorner\phi\urcorner = \ulcorner(t = s)\urcorner \wedge val(t) = val(s)).$$

Enthält L mehrere Basisprädikate, dann müssen mehrere Definitionen angegeben werden.

Hat man $Wahr_n$ einmal definiert, dann definiert man damit die Formel $Wahr_{n+1}(\ulcorner\phi\urcorner)$.

[1]Die Beweise und Definitionen dieses Abschnittes stammen alle aus [Halbach:1996].

$$Wahr_{n+1}(\ulcorner\phi\urcorner) =_{def}$$

$$Wahr_n(\ulcorner\phi\urcorner)\vee$$

$$(Satz(\ulcorner\phi\urcorner) \wedge kompl(\ulcorner\phi\urcorner) <= n+1\wedge$$

$$(\exists y((\ulcorner\phi\urcorner = (\ulcorner\neg\urcorner *_1 y)) \wedge Satz(y) \wedge kompl(y) <= n \wedge \neg Wahr_n(y))\vee$$

$$\exists y\exists z((\ulcorner\phi\urcorner = (\ulcorner(\urcorner *_1 y *_1 \ulcorner\vee\urcorner *_1 z *_1 \ulcorner)\urcorner)) \wedge (Wahr_n(y) \vee Wahr_n(z)))\vee$$

$$\exists y\exists z(Var(z) \wedge (x = (\ulcorner\forall\urcorner *_1 z *_1 y)) \wedge \forall w(Wahr_n(sub_1(y,z,w))))))$$

Für jede Menge bestimmter logischer Komplexität ist damit ein Wahrheitsprädikat definiert. Das Wahrheitsprädikat selbst kann eine andere logische Komplexität aufweisen, als die Komplexität der Ausdrücke für die es definiert wurde. Das Wahrheitsprädikat "$Wahr_0(\ulcorner\phi\urcorner)$" hat z.b. eine grössere Komplexität als ein Term.

Mit dieser Definition können wir jetzt zeigen, dass eine eingeschränkte Form der Tarski'schen Bikonditionale gültig ist.

Theorem 3.1.1 *Für jeden Satz ϕ aus L, der höchstens die Komplexität n hat gilt:*

$$\mathbf{PA} \vdash Wahr_n(\ulcorner\phi\urcorner) \leftrightarrow \phi$$

Aus beweistechnischen Gründen werden wir ein allgemeineres Theorem beweisen aus dem Theorem 3.1.1 folgt.

Theorem 3.1.2 *Für jede Formel $\phi(x_1,\ldots,x_n)$ aus L, die nur die angegebenen Variablen frei enthält und höchstens Komplexität n hat gilt:*

$$\mathbf{PA} \vdash Wahr_n(\ulcorner\phi(x_1,\ldots,x_n)\urcorner) \leftrightarrow \phi(x_1,\ldots,x_n)$$

Beweis 3.1.1 *Das Theorem wird bewiesen durch vollständige Induktion über den Index des Wahrheitsprädikats. $t_1(x_1,\ldots,x_n)$ ist ein Term, der höchstens die angegebenen Variablen frei enthält. Für alle Terme t gilt:*

$$\mathbf{PA} \vdash val(\ulcorner t(x_1,\ldots,x_n)\urcorner) = t(x_1,\ldots,x_n) \tag{1}$$

Mit Definition 3.1.2 und (1) folgt:

$$\mathbf{PA} \vdash Wahr_0(\ulcorner t_1(x_1,\ldots,x_n) = t_2(x_1,\ldots,x_n)\urcorner) \leftrightarrow$$

$$val(\ulcorner t_1(x_1,\ldots,x_n)\urcorner) = val(\ulcorner t_2(x_1,\ldots,x_n)\urcorner)) \leftrightarrow \qquad Def.\ 3.1.2$$

$$t_1(x_1,\ldots,x_n) = t_2(x_1,\ldots,x_n) \qquad mit\ (1)$$

Für $Wahr_0$ haben wir die Behauptung damit bewiesen. Unter der Annahme, die Behauptung gelte für $Wahr_n$ (Induktionsvoraussetzung I.V), beweisen wir dass sie für $Wahr_{n+1}$ gilt. Angenommen $\phi(x_1,\ldots,x_n)$ ist von der Form $\neg\psi(x_1,\ldots,x_n)$. Wenn ψ eine Komplexität kleiner gleich n hat, und man ausserdem noch folgende direkte Folgerung aus der Definition von "$Wahr_n$" verwendet,

$$\mathbf{PA} \vdash \forall x((Satz(x) \land kompl(\ulcorner\neg\urcorner *_1 x) <= n+1) \rightarrow \qquad (2)$$

$$(Wahr_{n+1}(\ulcorner\neg\urcorner *_1 x) \leftrightarrow \neg Wahr_n(x)))$$

dann kann man beweisen, dass:

$$\mathbf{PA} \vdash Wahr_{n+1}(\ulcorner\neg\psi(x_1,\ldots,x_n)\urcorner) \leftrightarrow$$

$$\neg Wahr_n(\ulcorner\psi(x_1,\ldots,x_n)\urcorner) \leftrightarrow \qquad mit\ (2)$$

$$\neg\psi(x_1,\ldots,x_n) \qquad I.V$$

Wenn $\phi(x_1,\ldots,x_n)$ von der Form $\psi(x_1,\ldots,x_n)\lor\chi(x_1,\ldots,x_n)$ ist, dann führt man den Beweis analog. Ist $\phi(x_1,\ldots,x_n)$ von der Form $\forall x_0\psi(x_0,x_1,\ldots,x_n)$, dann braucht man für den Beweis noch folgenden Satz:

$$\mathbf{PA} \vdash \forall x\forall y((Satz(\ulcorner\forall\urcorner *_1 y *_1 x) \land kompl(\ulcorner\forall\urcorner *_1 y *_1 x) <= n+1) \rightarrow \quad (3)$$

$$(Wahr_{n+1}(\ulcorner\forall\urcorner *_1 y *_1 x) \leftrightarrow \forall z Wahr_n(sub_1(x,y,z))))$$

D.h. wenn es wahr ist, dass alle natürllichen Zahlen ein Prädikat erfüllen, dann gilt für jede natürliche Zahl. dass sie dieses Prädikat erfüllt. Damit kann man jetzt den Beweis für Allsätze geben.

$$\mathbf{PA} \vdash Wahr_{n+1}(\ulcorner\forall x_0\psi(x_1,\ldots,x_n)\urcorner) \leftrightarrow$$

$$\forall w Wahr_n(\ulcorner\psi(w,x_1,\ldots,x_n)\urcorner) \leftrightarrow \qquad mit(3)$$

$$\forall w\psi(w,x_1,\ldots,x_n) \leftrightarrow \qquad I.V.$$

$$\forall x_0\psi(x_0,x_1,\ldots,x_n)$$

□

Für den Beweis des letztes Teils des Theorems haben wir die freien Variablen der Formeln verwendet, welhalb wir auch mit einer allgemeineren Form des Tarskischen T-Schemas gearbeitet haben.

Eine Tarski-Hierachie von Wahrheitsprädikaten liefert eine ähnliche Definition. Zuerst definieren wir Wahrheit für alle Sätze die kein Währheitsprädikat enthalten. Diese Definition nennen wir "$Wahr_0$". Dann definieren wir Wahrheit für alle Sätze die "$Wahr_0$" enthalten. Diese Definition nennen wir "$Wahr_1$" usw. Wahrheit für einen Satz, der Wahrheitsprädikate bis zur Stufe n enthält wird auf der Stufe $n+1$ definiert. Damit wird verhindert, dass "Der Lügner" und andere selbstreferentielle Sätze ausgewertet werden können. Die Komplexität der Sätze wird dann folgendermassen definiert:

Definition 3.1.3 *Wenn ϕ kein Wahrheitsprädikat enthält, dann ist:*

$$kompl_{\mathbf{T}}(\ulcorner \phi \urcorner) = 0$$

Wenn ϕ ein Wahrheitsprädikat mit Index n enthält, und n grösser oder gleich allen anderen Indices von Wahrheitsprädikaten ist, die ϕ enthält dann ist:

$$kompl_{\mathbf{T}}(\ulcorner \phi \urcorner) = n + 1$$

Dieser Komplexitätsbegriff ist implizit auch in unserer Definition enthalten. Denn ein Wahrheitsprädikat der Stufe $n+1$ ist komplexer als ein Wahrheitsprädikat der Stufe n. Ausserdem ist die Komplexität eines Wahrheitsprädikats der Stufe n immer grösser als n.

Diese Eigenschaften verhindern es, dass "$Wahr_n$" auf Sätze angewendet werden kann, in denen "$Wahr_n$" vorkommt. "Der Lügner" für das Wahrheitsprädikat der Stufe n kann nicht mit diesem Prädikat ausgewertet werden. Denn wenn die Komplexität von "$Wahr_n(\ulcorner \psi \urcorner)$" grösser als n ist, dann ist auch die Komplexität von "$\neg Wahr_n(\ulcorner \psi \urcorner)$" grösser als n.

Wir könnten allerdings versuchen eine Wahrheitsdefintion "$Wahr_n$" zu finden, die genau die Komplexität n hat, sodass "der Lügner" für dieses Prädikat die Komplexität $n+1$ hat. Das Wahrheitsprädikat wäre dann Teil der Klasse von Ausdrücken, deren Wahrheit es definiert.[2]

[2] Siehe [Kaye:1991], S 104.

3.2 Definierbarkeit der Wahrheit in der Logik 2.Stufe

Auf ein negatives Resultat dass wir bewiesen haben werden wir immer wieder zurückkommen. In einem formalen System, dass die Arithmetik enthält, können wir kein (vollständiges) Wahrheitsprädikat definieren.

Dieses Resultat kennzeichnet die Grenze dessen was wir definieren können und wird immer wieder verwendet um zu beweisen was nicht definiert werden kann.

Es scheint, als könnten wir die Menge der wahren Sätze der Arithmetik einfach angeben. Es sind einfach alle Sätze die im Standardmodell der Arithmetik wahr sind. Wollen wir diese naive Definition präzisieren, und die Definitionsmittel nutzen die uns die Arithmetik bietet (Codierung der eigenen Syntax) dann stossen wir auf Schwierigkeiten.

Jetzt könnten wir uns denken:"Na und. Wahrheit ist eben innerhalb der Arithmetik nicht definierbar. Hab ich ja schon immer vermutet." Was die Definition der Wahrheit innerhalb einer Theorie betrifft, sind wir damit in einen infinften Regress eingestiegen. Denn das Resultat gilt für jede Theorie die die Arithmetik enthält. Wenn wir arithmetische Wahrheit in einer stärkeren Theorie definieren, stellt sich dieselbe Frage wieder für diese Theorie.

Könnten wir nicht, ausgehend von unserer naiven Definition einen anderen Weg einschlagen? Vielleicht scheitert die Definition, weil wir die Sprachen in denen wir nach Definitionen suchen auf Sprachen in der Logik erster Stufe eingeschränkt haben. Versuchen wir die Ausdruckskraft der Arithmetik so zu erweitern, dass sie ihr eigenes Wahrheitsprädikat enthält. Aber nicht einfach dadurch, dass wir neue Prädikatsymbole hinzufügen (denn das bringt bezüglich Definierbarkeit der Wahrheit nichts), sondern indem wir unsere naive Definition ernst nehmen, und versuchen die Menge aller wahren Sätze (Sätze der Logik erster Stufe) der Arithmetik zu definieren. Diese Menge soll von jetzt an V heissen.[3]

3.2.1 Wie es geht

Die Erweiterung der Arithmetik, die ich untersuchen möchte, ist die Arithmetik zweiter Stufe. In dieser Theorie können wir nicht nur über Zahlen quantifizieren, sondern auch über Mengen von Zahlen. Die unendlich vielen

[3] V ist genaugenommen die Menge der Gödelnummern der wahren Sätze der Arithmetik.

Induktionsaxiome der Arithmetik, die wir aus dem Axiomenschema

$$(\phi(\ulcorner 0 \urcorner) \wedge \forall x(\phi(x) \rightarrow \phi(x+1))) \rightarrow \forall y(\phi(y))$$

erhalten, wenn wir für ϕ ein zahlentheoretisches Prädikat einsetzten, können wir in dieser erweiterten Theorie als ein Axiom formulieren.

$$\forall X((X(\ulcorner 0 \urcorner) \wedge \forall x(X(x) \rightarrow X(x+1))) \rightarrow \forall y(X(y)))$$

Die Variablen $X, Y, Z, \ldots$ quantifizieren dabei über Mengen von natürlichen Zahlen.

Eine Menge Σ ist in der Logik zweiter Stufe definierbar, wenn wir ein Prädikat P zweiter Stufe angeben können, sodass wenn $\sigma \in \Sigma$ dann P von σ erfüllt wird. Gegenüber dem bis jetzt verwendeten "syntaktischen" Definierbarkeitsbegriff, verwenden wir in diesem Abschnitt einen "semantischen" Definierbarkeitsbegriff. Dieser Schwenk ist notwendig, weil die Logik zweiter Stufe andere Eigenschaften besitzt (Unvollständigkeit) als die Logik erster Stufe.

Theorem 3.2.1 $\{V\}$ *ist in der Arithmetik erster Stufe definierbar.*

Um die Definierbarkeit einer Menge von Mengen zu beweisen, müssen wir die Sprache der Arithmetik erweitern. Nehmen wir zur Standardsprache zusätzlich ein einstelliges Prädikatsymobil G, dem wir immer die charakteristische Funktion einer Menge von natürlichen Zahlen zuordnen. [Boolos:1990]

Eine Menge von Mengen (kurz gesagt eine Klasse) M ist dann definierbar, wenn ein Satz F[4] angegeben werden kann, der G enthält, sodass:

$$g \in M \leftrightarrow F$$

z.b. ist die Klasse aller Mengen die 2 enthalten, genannt *Zwei* definierbar durch $G(2)$.

$$g \in Zwei \leftrightarrow G(2)$$

Wenn wir G immer die charakteristische Funktion von g zuordnen, dann haben wir eine Definition der Menge *Zwei*.

[4] F ist von der Logik erster Stufe her gesehen zwar ein Satz, enthält aber eine freie Variable die über Mengen quantifiziert, und ist daher von der Logik zweiter Stufe her gesehen kein Satz.

Beweis 3.2.1 *$\{V\}$ kann durch 5 Bestimmungen eindeutig charakterisiert werden.*

(1) V enthält nur Gödelnummern von Sätzen der Arithmetik. Mithilfe unseres Zusatzprädikats G formuliert:

$$\forall x(G(x) \to Satz(x))$$

(2) V enthält eine Gödelnummer eines Atomsatzes, wenn er in V_0 enthalten ist. (Weil jeder atomare Satz nur eine Identität zwischen Termen behauptet, können wir die Menge der wahren Atomsätze V_0 definieren.)[5]

$$\forall x(Satz_0(x) \to (G(x) \leftrightarrow V_0(x)))$$

(3)Die Gödelnummer von $\neg B$ ist in V enthalten, genau dann wenn die Gödelnummer von B nicht in V enthalten ist.[6]

$$\forall x \forall y(Satz(x) \land Negat(x,y) \to (G(y) \leftrightarrow \neg G(x)))$$

(4)Die Gödelnummer von $(B \lor C)$ ist in V, genau dann wenn entweder die Gödelnummer von B oder die Gödelnummer von C in V enthalten sind, oder beide.[7]

$$\forall x \forall y \forall z(Satz(x) \land Satz(y) \land Disjunktion(x,y,z) \to (G(z) \leftrightarrow (G(x) \lor G(y))))$$

(5)Für jede Formel F und Variable v gilt, dass wenn $\exists vF$ ein Satz ist, dann ist seine Gödelnummer in V, wenn es eine Zahl m gibt, sodass wenn man die freie Variable in $\exists vF$ durch den Standardnamen für m ersetzt die Nummer dieses Satzes in V ist.[8]

$$\forall x \forall y \forall z(Et(x,y,z) \land Satz(z) \to (G(z) \leftrightarrow \exists w \exists u(Si(x,y,w,u) \land G(u))))$$

[5] $Satz_0$ ist ein Prädikat, dass auf Gödelnummern von Sätzen ohne logische Zeichen zutrifft, also auf Atomsätze.

[6] $Negat(x,y)$ trifft zu, wenn y die Gödelnummer eines Satzes ist, der die Negation des Satzes ist, dessen Gödelnummer x ist.

[7] $Disjunktion(x,y,z)$ ist analog zu verstehen wie $Negat(x,y)$.

[8] $Et(x,y,z)$ trifft zu, wenn z die Nummer eines Satzes ist, der entsteht, wenn aus der Formel mit der Nummer x und der Variable mit der Nummer y eine existentielle Formel erzeugt wird. $Si(x,y,w,u)$ trifft zu, wenn u die Nummer einer Formel ist, die man erhält, wenn man in der Formel mit der Nummer x die Variable mit der Nummer y durch einen Standardnamen der auf w referiert ersetzt. Alle angegebenen Prädikate können in der Arithmetik definiert werden, unterscheiden sich aber leicht von den bisher verwendeten.

Wenn wir alle fünf Bestimmungen in einer Konjunktion zusammenfassen, dann erhalten wir eine Formel F, für die gilt:

$$g \in \{V\} \leftrightarrow F(G)$$

$\square$

Theorem 3.2.2 *V ist in der Arithmetik zweiter Stufe definierbar.*

Beweis 3.2.2 *Theorem 3.2.2 folgt aus Theorem 3.2.1. Wenn wir $\{V\}$ definiert haben durch einen Satz F, d.h.*

$$v \in \{V\} \leftrightarrow F$$

dann ersetzen wir jedes Vorkommen von G in F durch die einstellige Prädikatvariable X, und erhalten somit eine Formel $F(X)$ der Logik zweiter Stufe. Mittels dieser Formel können wir eine Formel

$$B(x) = \forall X(F(X) \rightarrow X(x))$$

bestimmen, die V definiert. Denn

$$k \in V \leftrightarrow$$

Für jede Menge A gilt, wenn G die char. Funktion von A zugewiesen wird, dass wenn $F(G)$ dann $G(k)$ $\leftrightarrow$

$$B(k)$$

$\square$

3.2.2 Warum (fast) niemand damit zufrieden ist

Eine Tatsache, die gegen die Akzeptanz der Logik zweiter Stufe spricht ist ihre Unvollständigkeit. Nicht alles was logisch aus einer Menge von Sätzen zweiter Stufe folgt, kann auch mittels logischer Schlussregeln aus dieser Satzmenge abgeleitet werden.

Ausserdem muss geklärt werden, wie die Quantoren der Logik zweiter Stufe zu interpretieren sind. Die Quantoren der Logik erster Stufe werden meist über die Terme interpretiert.

Ein Satz der Form $\forall x \Phi$ ist wahr genau dann wenn für alle Terme t gilt dass $\Phi(x/t)$ wahr ist. Der Allquantor wird definiert über die abzählbar unendliche Menge der Terme.

Da die Variablen der Logik zweiter Stufe über alle Teilmengen einer (unter Umständen) abzählbar unendlichen Menge laufen ist die Terminterpretation ausgeschlossen, wenn man nicht eine überabzählbare Menge von Termen annehmen will.

Dagegen lässt sich einwenden, dass auch in der Semantik der Logik erster Stufe schon überabzählbare Mengen verwendet werden um die Allgemeingültigkeit eines Quantorenschemas zu definieren.[9]

1.$\forall x(F(x) \rightarrow F(x))$ ist wahr genau dann wenn für alle offenen Sätze $\Phi(x)$ gilt, dass $\forall x(\Phi(x) \rightarrow \Phi(x))$ wahr ist.

Damit wird die Allgemeingültigkeit über die Menge der offenen Sätze definiert, und hängt damit von der Sprache ab, deren offene Sätze betrachtet werden.

Die Allgemeingültigkeit soll aber unabhängig von einer bestimmten Sprache sein. Will man sie so formulieren, dann bietet sich folgende Definition an.

2.$\forall x(F(x) \rightarrow F(x))$ ist wahr genau dann wenn für jede Extension H gilt, dass wenn F die Extension H hat, $\forall x(F(x) \rightarrow F(x))$ wahr ist.

Die Interpretation der Quantoren der Logik zweiter Stufe kann also gegeben werden, voraussgestzt die Allgemeingültigkeit von Quantorenschemata der Logik erster Stufe kann sprachunabhängig formuliert werden.

Die Unvollständigkeit der Logik zweiter Stufe lässt sich allerdings nicht so einfach an eine Eigenschaft der Logik erster Stufe binden. Die Vollständigkeit garantiert uns, dass alles was logisch aus einer Definition folgt, aus dieser Definition ableitbar ist.

Für manche Autoren sind aber gerade die Definitionsmöglichkeiten der Logik zweiter Stufe, und hier vor allem die Möglichkeit mathematische Theorien bis auf Isomorphie zu beschreiben ein grosser Vorteil dieser Logik. [Shapiro:1990]

In den weiteren Kapiteln werde ich mich nur mehr mit Theorien beschäftigen, die Wahrheitsprädikate im Rahmen der Logik erster Stufe definieren. Die Definitionsmöglichkeiten der Logik zweiter Stufe wurden nur vollständigkeitshalber angeführt.

[9]Siehe [Quine:1969], Seite 135.

Die Untersuchung des Lügnerparadoxons hat gezeigt, warum Wahrheitsprädikate nicht definiert werden können. Es hat sich gezeigt, dass die begrenzten Definitionsmögichkeiten der Sprache mit den unbegrenzten Ausdrucksmöglichkeiten der Sprache zusammenhängen.[10]

[10] Alle selbstreferentiellen Sätze können ausgedrückt werden.

Kapitel 4

Saul Kripke

4.1 Die Einwände gegen Tarski

Saul Kripke [Kripke:1984] hat eine Erweiterung der Tarskischen Wahrheitstheorie vorgeschlagen, die auf der Verwendung partieller Wahrheitsprädikate beruht. Er beginnt seine Untersuchung mit einer Analyse des Lügnerparadoxons. Wie viele andere Autoren misst Kripke den Erfolg seiner Theorie am Umgang mit dieser und anderen Paradoxien.

Das Lügnerparadxon wird in der Bibel erwähnt, es wird aber nicht als paradox erkannt. Vielmehr wird aus dem Paradox abgeleitet wie mit den Kretern umzugehen ist.

> "Es hat einer von ihnen (den Kretern) gesagt, ihr eigener Prophet :(*) **Die Kreter sind immer Lügner** , böse Tiere und faule Bäuche.
>
> Dieses Zeugnis ist wahr. Aus diesem Grund weise sie scharf zurecht, damit sie gesund werden im Glauben."[1]

Versuchen wir diese Aussage zu bewerten. Angenommen sie ist wahr, dann ist sie eine Lüge weil sie von einem Kreter geäussert wurde also falsch. Angenommen sie ist falsch. Aus dieser Annahme können wir nicht allgemien ableiten, dass die Aussage wahr sein muss. Es könnte ja sein, dass irdgendeine Aussage irgendeines Kreters wahr ist, und damit wäre (*) einfach falsch. In diesem allgemeinen Sinne handelt es sich also um keine Paradoxie.[2]

[1] Siehe [Bibel:1985], Titus 1, 12-13.

[2] Weil es sich bei (*) um keine echte Paradoxie handelt wird (*) oft in "Dieser Satz

Doch Kripke interessiert sich gerade deshalb für diese Aussage, weil sie nicht unter allen Umständen paradox ist, sondern nur wenn bestimmte Bedingungen erfüllt sind. Es muss der Fall sein, dass alle Aussagen von Kretern bis auf (*) Lügen sind. Wenn alle Aussagen der Kreter bis aus (*) Lügen sind, dann ist (*) genau dann wahr, wenn (*) falsch ist, also paradox.

Formal lässt sich diese Situation so beschreiben: Seien $P(x)$ und $Q(x)$ Prädikate von Sätzen. Dann kann es sein, aufgrund von empirischen Tatsachen, daß der Satz

$$(\forall x)(P(x) \rightarrow Q(x))$$

[oder $(\exists)(P(x) \wedge Q(x))$] selbst das Prädikat $P(x)$ erfüllt. Es kann auch vorkommen, daß dieser Satz als einziger $P(x)$ erfüllt. In diesem Fall sagt der Satz von sich selbst, daß er $Q(x)$ erfüllt. Ist $Q(x)$ das Prädikat "ist falsch", dann erhalten wir die Lügnerparadoxie. Für das obige Beispiel müsste man für $P(x)$ "wurde von einem Kreter behauptet" einsetzen.

Kripke versteht "ist wahr" als ein Prädikat, daß von Sätzen nicht von Propositionen oder Aussagen erfüllt wird. Kripke wählt Sätze als "truthbearer", weil für Sätze eine mathematische Theorie der Selbstreferenz entwickelt worden ist.

Warum für Propositionen bzw. Aussagen noch keine Theorie der Selbstreferenz entwickelt wurde, ist eine interessante Frage. Wenn wir "Proposition" als das verstehen was ausgesagt, geglaubt, gewusst, gehofft usw. werden kann, als etwas dass wahr oder falsch sein kann, dann ist es unmöglich für einen selbstreferentiellen und paradoxen Satz eine korrespondierende Proposition zu finden. Ein solcher Satz kann weder wahr noch falsch sein. Eine Lösung dieses Problems wäre, die selbstreferentiellen Propositionen als eine Teilmenge der selbstreferentiellen Sätze zu betrachten. Es würde einfach mehr selbstreferentielle Sätze als selbstreferentielle Propositionen geben. In Bezug auf eine Lösung des Lügnerparadoxons ist das aber unbefriedigend. Der Lügner wäre kein Problem für eine Wahrheitstheorie, weil er es nicht schaftt eine Proposition auszudrücken. Wenn wir unter einer Aussage etwas verstehen, dass ausgesagt werden kann und wahr oder falsch sein kann, dann existiert dasselbe Problem auch in Bezug auf Aussagen. Das ist bemerkenswert, weil Aussagen etwas konkreteres als Propositionen zu sein scheinen.

---

ist falsch" umformuliert. Dabei handelt es sich in jedem Fall um eine Paradoxie, die von keinen emprischen Bedingungen abhängt.

Ein anderes Problem besteht darin, Identitätsbedingungen für Propositionen anzugeben. Bei der obigen Bestimmung der Proposition habe ich die Identitätsbedingungen schon vorausgestzt und Proposition nur allgemein charakterisiert.

Es sollte erwähnt werden, dass einige Autoren Probleme haben Identitätsbedingungen für Propositionen bzw. Aussagen anzuerkennen. Zusammen mit dem Grundsatz "No entity without identity" [Quine:, S.206] folgt daraus dass es solche Entitäten nicht gibt. [Quine:, S.112]

Die Versionen des Lügners, die empirische Prädikate verwenden, verweisen für Kripke auf einen zentralen Punkt.

> "...many, probably most, of our ordinary assertions about truth and falsity are liable, if the empirical facts are extremely unfavorable, to exhibit paradoxical features."[3]

In dieser Behauptung manifestiert sich eine spezielle Art von Skeptizismus. Die Wahrheit der Sätze die wir äussern liegt ausserhalb unserer Kontrolle. Wir können uns noch so bemühen immer die Wahrheit zu sagen, dadurch dass andere etwas über uns und unsere Aussagen sagen können oder durch andere empirische Fakten entstehen vielleicht paradoxe Konstellationen.

Beim obigen Beispiel ist das ungünstige empirische Faktum die Tatsache, dass alle Aussagen von Kretern bis auf (*) Lügen sind. Wäre das nicht so, dann könnte (*) zwar nicht wahr aber falsch sein.

Zur obigen Behauptung bringt Kripke ein Beispiel, das ich in leicht abgewandelter Form wiedergebe. Angenommen ich behaupte dass:

> (1)*Die meisten (mehr als die Hälfte) von Hans Dichands Behauptungen über Politik sind falsch.*

Nichts an (1) ist wesentlich falsch, d.h. (1) ist nicht notwendigerweise falsch, es wäre zumindest möglich, das (1) wahr ist. Es gibt also kein semantisches Kriterium an dem erkennbar wäre, das (1) notwendigerweise falsch ist. Um herauszufinden ob (1) wahr ist, könnte ich alle Aussagen Dichands über Politik durchnumerieren, und bewerten. Was wäre aber, wenn Dichand ebenso viele wahre wie falsche Aussagen über Politik gemacht hat (So unwahrscheinlich das auch ist!), bis auf eine proplematische Aussage.

[3]Siehe [Kripke:1984], S 54.

(2)*Alles was Michael Pucher über Politik sagt ist wahr.*[4]

Wenn (1) meine einzige Aussage über Politik ist, oder alle meine anderen Aussagen über Politik wahr sind, dann ist die Konjunktion aus (1) und (2) ein Paradox.

Angenommen die Konjunktion aus (1) und (2) wäre wahr. Dann wäre (2) wahr. Dann müsste aber (1) falsch sein, weil die Falschheit von (2) die Bedingung der Wahrheit von (1) ist. Also wäre die Konjunktion falsch.

Angenommen die Konjunktion wäre falsch. Dann wäre entweder (1) oder (2) falsch. Wenn (1) falsch ist, dann muss (2) wahr sein. Dann muss aber (1) als eine Aussage von mir wahr sein. Wenn (2) falsch ist, dann ist (1) wahr. Weil (1) aber meine Aussage ist muss sie falsch sein damit (2) falsch ist. Wie oben muss dann aber (2) wahr sein.

Dass die Konjuktion aus (1) und (2) paradox ist, muss weder mir noch Hans Dichand bewusst sein.

Um zu illustrieren dass nicht einmal dem besten Paradoxieexperten jede Paradoxie bewusst ist, erzählt Kripke eine Geschichte über Bertrand Russell und George Edward Moore in der Russell eine Paradoxie, die sich in einer Diskussion ergeben hat unbewusst geblieben ist. Russell fragte Moore, ob dieser immer die Wahrheit gesagt habe, und bewertete Moores negative Antwort als die einzige falsche Aussage Moores. Wenn Russell dachte, dass alle anderen Aussagen Moore's wahr seien, dann war die Antwort Moore's nicht nur falsch, sondern paradox.

> "Die Moral: Eine adequate Theorie muss es unseren Aussagen, die den Begriff der Wahrheit enthalten, erlauben *riskant* zu sein. Sie riskieren paradox zu sein..."

Nach diesen Beobachtungen beginnt Kripke seine Theorie zu entwickeln. Ein zentraler Begriff von Kripkes Theorie ist der Begriff der "Gründung" bzw. des "gegründeten Satzes".[5]
Wenn ein Satz das Prädikat "ist wahr" enthält, hängt seine Wahrheit von anderen Sätzen ab, die zuerst bewertet werden müssen. Wenn dieser Prozess terminiert, dann ist der ursprüngliche Satz gegründet (grounded), sonst ist

[4]Ich nehme an, dass eine Aussage darüber was jemand über Politik sagt, auch eine Aussage über Politik ist.

[5]Dieser Begriff wurde wahrscheinlich das erste Mal von Hans Herzberger in [Herzberger:1970] explizit verwendet.

er ungegründet (ungrounded).
Nicht jeder ungegründete Satz ist paradox.

(3) (3) ist wahr.

(3), der sogenannte Wahrsager (truthteller) ist ungegründet, aber nicht paradox. Er kann etweder wahr oder falsch sein, je nachdem wofür ich mich entscheide. Den Zusammenhang zwischen dem Begriff der Gegründetheit und dem vorher formulierten Skeptizismus erklärt Kripke so:

> "We make utterances which we hope will turn out to be grounded." [6]

Wir können dem Skeptizismus entgehen, wenn wir uns vergewisssern, dass unsere Aussagen (Sätze) gegründet sind.[7]

Kripke bringt drei Einwände gegen die Tarskische Wahrheitstheorie vor und zeigt dass sie nicht adequat mit den undegründeten Sätzen umgehen kann.

Die Tarskische Wahrheitstheorie ermöglicht die Konstruktion von unendlich vielen Wahrheitsprädikaten.

Sei L_0 eine formale Sprache, in der alle Prädikate total definiert sind, dann können wir eine Sprache L_1 definieren, die ein Wahrheitsprädikat für L_0 enthält. Dieser Prozeß kann wiederholt werden, sodaß wir eine Hierachie von Sprachen erzeugen $\{L_0 L_1 L_2 ...\}$ wobei L_{i+1} ein Wahrheitsprädikat für L_i enthält.[8]

Erster Einwand Wenn ich versuche meine Behauptung über Hans Dichand zu verifizieren, weiss ich nicht welches Wahrheitsprädikat ich verwenden soll, da ich die höchste Ebene von Dichands Aussagen nicht kenne. Kripke meint, es sei eher so, dass die Aussage ihre eigene Ebene sucht, auf der sie bewertet werden muss.

Und selbst wenn ich die Ebene der Aussage kennen würde, könnte ich nicht sicher sein dass die Aussage immer auf dieser Ebene bewertet werden

[6] [Kripke:1984], S 57

[7] Am Ende der Darstellung von Kripkes Theorie wird sich herausstellen, dass es nicht so einfach ist.

[8] Ein bekannter Einwand gegen diese Vorgehensweise ist, dass unsere Sprache nur ein Wahrheitsprädikat enthält, und nicht unendlich viele. Akzeptiert man diesen Einwand, dann kann man entweder die Untersuchung natürlicher Sprachen völlig ausklammern, oder zu beweisen versuchen dass die verschiedenen Ebenen auch in der natürlichen Sprache vorhanden sind.

kann. Die Ebene der Aussage ist nicht fix wie in der Tarski-Hierachie, sondern verändert sich mit den Aussagen Hans Dichands, also mit den empirischen Fakten.

Zweiter Einwand Angenommen ich behaupte dass:

(4)*Alles was Dichand über Politk sagt ist falsch.*

und Hans Dichand behauptet:

(5)*Alles was Michael Pucher über Politk sagt ist falsch.*

Dann möchte ich, dass meine Behauptung auch (5) einschließt. Das ist aber in der Theorie, in der jeder Satz eine Ebene besitzt, die für ihn **wesentlich** ist nicht möglich. Trotzdem kann ich mir Situationen vorstellen, in denen (4) und (5) bewertet werden können.
Angenommen ich hätte etwas wahres über Politik gesagt, dann wäre (5) falsch, auch ohne Bewertung von (4). Sind auch alle anderen Aussagen Dichands über Politik falsch, dann ist (4) wahr, hat Dichand etwas wahres über Politik gesagt, dann ist (4) falsch. Zu bemerken ist, daß im letzeren Fall (5) nicht vorher bewertet werden muß, während im ersten Fall (5) zuerst bewertet werden muß.

Dritter Einwand Für die Tarskische Theorie ist es keine Schwierigkeit Sätze wie "Schnee ist weiß", ""Schnee ist weiß" ist wahr" usw. zu bewerten. Aber die Aussage, daß alle Sätze dieser Sequenz wahr sind bereitet Schwierigkeiten.

4.2 Die Lösung

Um diesen Einwänden zu begegnen, muss ein Zusammenhang zwischen der Bewertungsebene eines Satzes und den empirischen Fakten hergestellt werden. Die Ebene auf der ein Satz bewertet werden muss, hängt nicht nur von seinen syntaktischen Eigenschaften ab[9], sondern auch von empirischen Tatsachen, wie der Tatsache dass Hans Dichand dieses und jenes gesagt hat.

Nach meiner Ansicht sind alle drei Einwände berechtigt, wobei der zweite und der dritte Einwand am stärksten sind. Der erste Einwand könnte durch

[9]Wenn er höchstens Wahrheitsprädikate vom Typ i enthält, dann muss er durch das Wahrheitsprädikat T_{i+1} bewertet werden.

den Hinweis geschwächt werden, dass die Tarski-Hierachie nicht voraussetzt dass ich weiss auf welcher Ebene ein Satz zu bewerten ist. Dann bleibt noch immer das Problem der wesentlichen Bewertungsebene.

Der zweite und der dritte Einwand zeigen dass bestimmte Sätze, die intuitiv verständlich sind und bewertet werden können, in der Tarski-Hierachie keinen Platz finden.

Kripke verwendet für seine Wahrheitstheorie Sprachen, die partiell definierte Prädikate enthalten können. Wie Strawson[10] versteht Kripke einen Satz als einen Versuch eine Proposition auszudrücken. Die Wohlgeformtheit, bzw. der Sinn des Satzes liegt in der Tatsache das es spezifizierbare Umstände gibt, unter denen der Satz eine Proposition ausdrückt. So ist (1) zwar immer sinnvoll (meaningful), aber manchmal drückt er keine Proposition aus.

Wenn man Sätze so versteht kann man eine mehrwertige Logik verwenden, ohne den zusätzlichen Wert als eigenen Wahrheitswert aufzufassen. Ein Satz hat diesen zusätzlichen Wert, wenn er keine Proposition ausdrückt. Nur Sätze die entweder wahr oder falsch sind drücken Propositionen aus. Die anderen Sätze haben keinen Wahrheitswert und sind nur als solche gekennzeichnet die keine Proposition ausdrücken.

Die Semantik dieser mehrwertigen Logik ist folgendermassen definiert. Die Semantik der partiellen Prädikate wird folgendermassen definiert. Sei D ein nichtleeres Diskursuniversum, ein monadisches Prädikat $P(x)$ wird interpretiert durch ein Paar (S_1, S_2) von disjunkten Teilmengen von D. Die Vereinigung von S_1 und S_2 muss nicht gleich D sein. Darum kann das Paar (S_1, S_2) als "Extension" eines partiellen Prädikats interpretiert werden. S_1 ist die Extension von $P(x)$, S_2 die Anti-Extension. S_1 enthält die Elemente von D die $P(x)$ erfüllen, S_2 enthält die Elemente von D die $\neg P(x)$ erfüllen.

Die Bewertungsregeln sind jene von S.C.Kleene's [Kleene:1991] starker dreiwertiger Logik.[11]

[10]Kripke interpretiert Strawson als würde er behaupten, dass "the present king of france is bald“ scheitert eine Proposition auszudrücken, aber trotzdem sinnvoll ist, weil er Richtungen (Bedingungen) angibt eine Proposition auszudrücken.

[11]Kleene's Begriff der regulären Tabellen ist äquivalent zur Forderung der Monotonie von Φ.(siehe unten) Kripke wundert sich das seine Verwendung der dreiwertigen Logik von Kleene manchmal als Beispiel angeführt wird für eine Aufgabe der Standardlogik, oder die Annahme eines neuen Wahrheitswertes."Undefiniert“ ist für Kripke wie oben schon bemerkt kein eigener Wahrheitswert. Konventionen um mit Sätzen umzugehen die keine Proposition ausdrücken, betrachtet er nicht als signifikante Veränderungen der Logik. Ausserdem kann Kripke's gesamte Theorie in einer klassischen Metasprache formalisiert

Anhand folgender fiktiver Lernsituation veranschaulicht Kripke den Prozess der Gründung eines Satzes.

Angenommen wir wollen jemandem die Verwendung des Wortes "wahr" erklären. Zuerst werden wir ihm erklären, dass "wahr" ein einstelliges Prädikat ist welches von Sätzen erfüllt wird. Dann werden wir Bewertungsregeln angeben mit denen entschieden werden kann ob ein Satz dieses Prädikat erfüllt oder nicht.

Ich nenne diese Situation fiktiv, weil es fraglich ist ob jemand verstehen kann was ein Satz ist ohne zu verstehen was "wahr" bedeutet. Eine Bedingung um "wahr" anwenden zu können ist aber zu verstehen was ein Satz ist. Die Situation wäre realistischer, wenn man das Prädikat "uhu" nennen würde und der (die) Lernende nach Angabe der Regeln erkennen würde dass "uhu" äquivalent zu "wahr" ist. Eine andere realistischere Situation wäre jemandem der nicht deutsch spricht die Verwendung des Wortes "wahr" zu erklären, wobei sie (er) nachdem er (sie) die Verwendung des Wortes verstanden hat erkennen würde, dass (sie) er ein gleichbedeutendes Wort in seiner (ihrer) Sprache kennt.

Die Bewertungsreglen könnten so aussehen.

> Du hast das Recht von einem Satz zu behaupten dass er wahr ist, wenn du das Recht hast den Satz zu behaupten.
>
> Du hast das Recht von einem Satz zu verneinen dass er wahr ist, wenn du das Recht hast den Satz zu verneinen.

Damit hat der (die) Lernende vielleicht noch Schwierigkeiten Sätze zu bewerten die das Wort "wahr" enthalten.

Wie man sich diese Bewertung veranschaulichen kann, zeigt folgendes Beispiel. Angenommen wir wollen folgenden Satz bewerten.

> (7)*Einige Sätze die in der Kronenzeitung vom 7. Oktober, 1971 stehen sind wahr.*

Wenn der Wahrheitswert von (7) unklar ist, so ist auch der Wahrheitswert von (8) unklar.

> (8) (7) ist wahr.

---

werden.

Angenommen einer der Sätze der besagten Ausgabe der Kronenzeitung war

(6) Schnee ist weiss

Wenn nun jemand bereit ist (6) zu behaupten, dann ist er nach obiger Regel auch bereit "(6) ist wahr" zu behaupten. Daraus folgt aber mittels existentieller Generalisierung[12] (7), und daraus folgt mittels unserer Bewertungsreglen (8).

Solche Sätze wie (6), (7) und (8) werden in diesem Prozess also bewertet werden. Ein ungegründeter Satz wie (3) wird durch diesen Prozess keinen Wahrheitswert erhalten. Ein Satz ist ungegründet, wenn er in diesem Prozess keinen Wahrheitswert erhält.

Dieser Prozess kann auch in der Modelltheorie formalisiert werden. Sei L eine Sprache erster Stufe, mit endlich oder abzählbar unendlich vielen Prädikaten. Die Variablen von L werden über einem Wertebereich D interpretiert, und die Prädikate sind für jedes Element aus D definiert, also total. L sei reich genug um seine eigene Syntax auszudrücken, und es existiere ein Kodierungsschema das endliche Tupeln aus D in Elemente aus D kodiert. Wie wir gesehen haben reicht es aus dass L die Sprache der Arithmetik enthält um diese Bedingungen zu erfüllen.

Zu L wird ein monadisches Prädikat $T(x)$ hinzugefügt, dass nur partiell definiert sein muß. Die neue Sprache heisse $\mathcal{L}$.

Sei $\mathcal{L}(S_1, S_2)$ die Interpretation von $\mathcal{L}$ die entsteht wenn $T(x)$ durch (S_1, S_2) interpretiert wird. S_1 ist die Extension, S_2 die Antiextension von $T(x)$.

Sei S_1' die Menge von Codes von wahren Sätzen von $\mathcal{L}(S_1, S_2)$, S_2' die Menge von Elementen von D welche entweder keine Codes von Sätzen sind, oder falsche Sätze von $\mathcal{L}(S_1, S_2)$ sind. Ein Paar (S_1, S_2) welches die Bedingung erfüllt, dass $S_1 = S_1'$ und $S_2 = S_2'$ ist ein Fixpunkt.

Die Funktion Φ sei folgendermassen definiert.

$$\Phi : \mathfrak{P}(D) \times \mathfrak{P}(D) \to \mathfrak{P}(D)$$
$$\Phi(S_1, S_2) = (S_1', S_2')$$

[12] Die Regel der existentiellen Generalisierung bzw. Regel zur Einführung eines Existenzquantors besagt, dass wenn ein bestimmter Gegenstand eine Eigenschaft hat es etwas gibt dass diese Eigenschaft hat.

Ein Fixpunkt von Φ ist ein Paar, welches unter Φ obige Bedingung erfüllt. Die Existenz von Fixpunkten von Φ soll bewiesen werden.

Wie sollen diese Definitionen den Prozess der Bewertung wiederspiegeln? Angenommen (S_1, S_2) ist eine Interpretation von $T(x)$. Wenn wir jetzt die Funktion Φ auf diese Menge anwenden, wird sie entweder grösser werden oder gleich gross bleiben (In diesem Fall hätten wir einen Fixpunkt erreicht.). Warum ? Sei $\ulcorner\phi\urcorner \in S_1$. Dann wird unter dieser Interpretation "$T(\ulcorner\phi\urcorner)$" wahr sein. Der Code dieses Satzes ist nicht notwendigerweise in S_1 enthalten, aber in S_1'. Analog für S_2.

Um die Existenz von Fixpunkten zu beweisen definieren wir $\mathcal{L}_0$ als $\mathcal{L}(\emptyset, \emptyset)$. Angenommen für eine natürliche Zahl α sei $\mathcal{L}_\alpha$ schon definiert, dann sei

$$\mathcal{L}_{\alpha+1} = \mathcal{L}(S_1', S_2').$$

Die so erzeugte Hierachie ist analog zur Tarskischen Hierachie, denn $T(x)$ interpretiert in $\mathcal{L}_{\alpha+1}$ ist ein Wahrheitsprädikat für $\mathcal{L}_\alpha$. Nach dieser Definition ist Φ ein monotoner Operator. Wenn

$$(S_1, S_2) \leq (S_1^\dagger, S_2^\dagger),$$

dann ist

$$\Phi((S_1, S_2)) \leq \Phi((S_1^\dagger, S_2^\dagger)).$$

Ein Satz der einmal bewertet wurde, behält diesen Wert in jeder höheren Ebene. Nun kann für jedes α bewiesen werden, dass jede Interpretation von $T(x)$ in $L_{\alpha+1}$ die Interpretation von $T(x)$ in L_α erweitert.

Bisher haben wir für jede natürliche Zahl eine Ebene definiert. Die Erweiterung für transfinite Ordinalzahlen sieht folgendermassen aus.
Sei

$$L_\omega = \mathcal{L}(S_{1,\omega}, S_{2,\omega})$$

wobei

$$S_{1,\alpha} = \bigcup_{\beta<\alpha} S_{1,\beta}$$

Danach können $L_{\omega+1}, L_{\omega+2}$ usw. definiert werden.

Formal definieren wir die interpretierten Sprachen L_α für beliebige Ordinalzahlen α so: Wenn α ein Nachfolger ist ($\alpha = \beta + 1$), dann ist

$$L_\alpha = \mathcal{L}(S_{1,\alpha}, S_{2,\alpha})$$

wobei $S_{1,\alpha}$ die Menge von Codes von wahren Sätzen von L_β ist, und $S_{2,\alpha}$ die Menge aller Elemente von D, die Codes von falschen Sätzen von L_β sind, oder nicht Codes von Sätzen von L_β sind.

Ist α ein Grenzwert, dann ist $S_{1,\alpha} = \bigcup_{\beta<\alpha} S_{1,\beta}$. Auch bei transfiniten Ebenen ist es wahr, dass $T(x)$ größer wird (oder gleich groß bleibt), wenn α größer wird.

Die Existenz eines Fixpunktes von Φ ergibt sich aus einer Überlegung bezüglich der Anzahl der Sätze. Da die Menge aller Sätze von $\mathcal{L}$ eine Menge bildet die eine bestimmte Kardinalität (Anzahl) besitzt, muss der Prozess irgendwann stoppen, das erste Mal beim kleinsten Fixpunkt von Φ. In jedem grösseren Fixpunkt bleiben diejenigen Sätze wahr, die im kleinsten Fixpunkt wahr sind. Wie ich im Abschnitt über Fixpunkte gezeigt habe, kann die Existenz von Fixpunkten auch anders bewiesen werden.

K. ist der Ansicht, dass die formale Konstruktion der Fixpunkte dem was jemand macht der lernt das Wahrheitsprädikat zu verwenden sehr nahe kommt.

K. überlegt und skizziert auch die Konstruktion einer Sprache, die ihr eigenes Erfüllungsprädikat enthält. Das Wahrheitsprädikat so wie es eben definiert wurde, kann dazu verwendet werden. Dazu muss L einen Namen für jedes Objekt seines Diskursuniversums besitzen, und eine Denotationsrelation definiert werden. Dann kann der Begriff der Erfüllung durch den Begriff der Wahrheit definiert werden. Anstatt zu sagen, dass $A(x)$ von einem Gegenstand a erfüllt wird, sagen wir einfach, dass $A(x)$ wahr wird, wenn die Variable durch einen Namen von a ersetzt wird.
Ist in L kein Name für jedes Objekt verfügbar, dann muss ein anderer Weg eingeschlagen werden. L kann zu $\mathcal{L}$ erweitert werden, durch Hinzunahme eines Erfüllbarkeitsprädikates $Sat(s, x)$, wobei s über endliche Tupeln von Elementen von D läuft, und x über Formeln von D. Analog der Konstruktion des Wahrheitsprädikats können wir so eine Hierachie von Sprachen definieren und erreichen schließlich einen Fixpunkt, eine Sprache die ihr eigenes Erfüllbarkeitsprädikat einhält. Wenn L abzählbar ist, D aber nicht, so schließt die Konstruktion des Wahrheitsprädikates bei einer abzählbaren Ordinalzahl, die Konstruktion des Erfüllbarkeitsprädikates aber eventuell bei einer überabzählbaren Ordinalzahl. Der Grund dafür liegt in der Definition von

$Sat(s, x)$.

K. hält das Erfüllbarkeitsprädikat für grundlegender als das Wahrheitsprädikat, weil es einfacher definiert werden kann. Nachdem aber gezeigt ist, dass eine Definition des einen durch das andere möglich ist, kann man sich auch mit dem weniger grundlegenden beschäftigen.

Eine einfachere Definition des Erfüllbarkeitsprädikats könnte so aussehen. Angenommen L besitzt einen kanonischen Namen für jedes Element aus D. Dann können wir Paare (A, T), (A, F) betrachten, sodass A wahr, bzw. falsch ist. Die Kleene Regeln entsprechen Abschlussbedingungen für solche Paare, d.h wenn $(A(a), F) \in S$ für jeden Namen a, dann ist $((\exists)A(x), F) \in S$. Wenn $(A(a), T) \in S$ dann ist $((\exists)A(x), T) \in S$. Angenommen k sei ein Name von A, und wenn $(A, T) \in S$, dann sei $(T(k), T) \in S$, wenn $(A, F) \in S$, dann sei $(T(k), F) \in S$. Eine Menge S, die unter diesen Bedingungen abgeschlossen ist, ist ein Fixpunkt bezüglich Wahrheit, d.h. mit einem Erfüllbarkeitsprädikat kann ein Wahrheitsprädikat definiert werden. K. meint, diese Definition sei einfacher als jene im Text, und er habe sie verwendet um zu zeigen, dass der Fixpunkt induktiv in Moschovakis Sinn ist, sie erlaubt aber keine so einfache Paralellisierung zu Tarski, da der Begriff der Ebene fehlt.
Die Konstruktion eines Erfüllbarkeitsprädikats könnte auch für eine Sprache durchgeführt werden, die wesentlich reicher ist als L, z.b. könnte sie einen Quantor "für überabzählbar viele x" enthalten.

Erfüllt die Kripke'sche Theorie die Bedingungen die Kripke selbst an eine Wahrheitstheorie gestellt hat?

Begegnung des ersten Einwands - Kenntnis der Ebene eines Satzes wird vorausgesetzt.

In Kripkes Modell wird die Kenntnis der Ebene eines Satzes nicht vorausgesetzt, weil ein Satz gar keine Ebene besitzt auf der er immer bewertet werden muss. Die Bewertungsebene eines Satzes hängt von den empirischen Tatsachen ab.

Ich möchte dass an einem Beispiel illustrieren. Vorher muss aber noch etwas zum Begriff der Gegründetheit gesagt werden.

Der Begriff der Gegründetheit kann so definiert werden: Ein Satz A ist gegründet, wenn A einen Wahrheitswert im kleinsten Fixpunkt $\mathcal{L}_\sigma$ besitzt. Die Ebene von A wird als kleinste Ordinalzahl α definiert in der A einen Wahrheitswert besitzt.

Da $\mathcal{L}$ die Zahlentheorie enthält, kann der Lügner bzw. der Wahrsager in $\mathcal{L}$ ausgedrückt werden. Es ist leicht zu zeigen, dass beide Sätze ungegründet sind. Die Tatsache, dass es für gewöhnliche Sätze keine Garantie gibt, dass sie gegründet sein werden und dass ihre Ebene nicht im vorhinein bekannt ist, zeigt sich sehr schön in Kripkes Modell.

Sei $T(x)$ das Wahrheitsprädikat von $\mathcal{L}$.

$$(9)(\forall x)((P(x) \rightarrow \neg T(x))$$

Dieser Satz muss in unterschiedlichen Kontexten auf verschiedenen Ebenen bewertet werden. Angenommen $P(x)$ trifft nur auf (9) selbst zu, dann wird (9) der Satz sein, der von sich selbst behauptet falsch zu sein und ungegründet sein. Wenn $P(x)$ auf mehrere Sätze der Ebene 2,4 und 13 zutrifft, dann wird (9) auf der Ebene 14 begründet sein. Trifft $P(x)$ auf beliebige Sätze endlicher Ebenen zu, dann wird (9) auf der Ebene ω gegründet sein.

$P(x)$ kann entweder ein syntaktisches Prädikat sein, oder sich auf Aussagen beziehen. Es ist daher nicht von vornherein klar auf welche und wieviele Sätze sich $P(x)$ bezieht. Es sieht so aus, als ob diese verschiedenen Ebenen der Bewertung dadurch zustande kommen, weil $P(x)$ nicht genauer spezifiziert wird. Bei einer genaueren Spezifikation von $P(x)$ sollte es nur eine mögliche Bewertungsebene geben, und damit wäre auch der erste Einwand gegenüber Kripkes Theorie gerechtfertigt.

Dieser Einwand ist nicht gerechtfertigt, ausser man verlangt von einer genaueren Spezifikation die Angabe des gesamten Modells in dem der Satz bewertet werden soll, womit aber die Rede von verschiedenen Interpretationen eines Satzes sinnlos wäre.

Bezieht sich $P(x)$ auf Aussagen, dann kann das Prädikat noch so genau spezifiziert sein, die Bewertungsebene ergibt sich doch erst aus der Kenntnis der Aussagen auf die sich $P(x)$ bezieht.

Begegnung des zweiten Einwands - (4) und (5) können nicht bewertet werden.

Betrachten wir noch einmal die Sätze (4) und (5). Wir können (4) als (9) formalisieren, indem wir $P(x)$ als x ist ein von Hans Dichand geäusserter Satz zur Politik" interpretieren. (5) können wir als

$$(10)(\forall x)((Q(x) \rightarrow \neg T(x))$$

interpretieren, wobei $Q(x)$ x ist ein von Michael Pucher geäusserter Satz zur Politik" bedeutet. (10) befindet sich in der Extension von $P(x)$ und (9)

in der Extension von $Q(x)$. Nichts garantiert, dass (9) und (10) gegründet sein werden. Angenommen es gibt einen Satz der Ebene α der $Q(x)$ erfüllt. Dann wird (10) falsch und gegründet sein auf der Ebene $\alpha + 1$. Wenn auch alle anderen Sätze die $P(x)$ erfüllen falsch sind, dann wird (9) wahr sein und gegründet. Die Ebenen von (9) und (10) beruhen also auf empirischen Tatsachen.
Es wurde erwähnt, dass Satze wie (3), obwohl sie ungegründet sind, nicht paradox sind. Wie kann das in dieser formalen Theorie dargestellt werden. Sei (3) durch

$$(11)(\forall x)((Q(x) \rightarrow T(x))$$

formalisiert, wobei $Q(x)$ ein nur von (11) (d.h. von der Gödelnummer von (11)) erfüllt wird. Angenommen wir definieren zu Beginn als das Wahrheitsprädikat noch undefiniert ist, dass (11) wahr ist. (11) wird auf allen Ebenen wahr bleiben, da $Q(x)$ ja nur auf (11) zutrifft. Der Wahrheitswert von (11) kann also beliebig festgestzt werden.

Begegnung des dritten Einwands - Sätze in denen über "Wahrheit" quantifiziert wird können nicht bewertet werden. Die obigen Beispiele zeigen, dass es kein prinzipielles Problem ist über Wahrheit" zu quantifizieren. Durch die Verwendung einer mehrwertigen Logik entsteht auch kein Problem wenn man über ungegründete Sätze quantifiziert. Eine (positive) Allquantifikation über einen ungegründeten Satz wird ebenfalls ungegründet sein. Bei einer (positiven) Existenzquantifikation genügt es, wenn ein Satz die Bedingung der Quantifikation erfüllt.

Damit ist gezeigt, dass Kripkes Theorie allen Einwänden die Kripke gegen Tarskis Theorie vorbringt begegnen kann.

Ich möchte jetzt noch skizzieren, wie Kripke innerhalb seiner Theorie mit Paradoxien umgeht, und seinen Vergleich mit der Tarskischen Theorie herausarbeiten.

Ein Satz ist paradox, wenn er in keinem Fixpunkt einen Wahrheitswert hat (wenn er keinen Wahrheitswert in irgendeinem Fixpunkt hat).

$$A \textit{ ist paradox} \leftrightarrow (\Phi((S_1, S_2)) = (S_1, S_2) \rightarrow (\neg A \in S_1 \wedge \neg A \in S_2))$$

(9) ist paradox, falls $P(x)$ nur auf (9) selbst zutrifft.[13] Ist dieses Zutreffen eine empirische Tatsache, ist (9) aus empirischen Gründen ein Paradox.
Unter Verwendung des Zornschen Lemmas kann man zeigen, dass jeder Fixpunkt zu einem maximalen Fixpunkt erweitert werden kann.[14]
Es gibt auch ungegründete, nicht paradoxe Sätze, die in jedem Fixpunkt denselben Wahrheitswert haben wenn sie einen Wahrheitswert haben.

(12)Entweder (12) oder seine Negation ist wahr.

Da ein Satz wahr ist, wenn eines seiner Disjunkte wahr ist, ist (12) wahr, wenn er als wahr initialisiert wird, wennn er als falsch initialisiert wird, ist die Negation von (12) wahr, somit beide Konjunkte, somit (12) selbst. Es gibt Fixpunkte, in denen (12) wahr ist, aber keine in denen (12) falsch ist.

Ein Fixpunkt wird wesentlich genannt (intrinsisch), wenn er keinem Satz einen Wahrheitswert zuordnet, der mit seinem Wahrheitswert in einem anderen Fixpunkt in Konflikt ist. D.h. (S_1, S_2) ist ein wesentlicher Fixpunkt, genaudann wenn es keinen anderen Fixpunkt $(S_1^\dagger, S_2^\dagger)$ und Satz A aus L' gibt, sodass $A \in (S_1 \cap S_2^\dagger) \cup (S_2 \cap S_1^\dagger)$.$A$ besitzt einen wesentlichen Wahrheitswert, wenn er einen Wahrheitswert in einem wesentlichen Fixpunkt erhält. (12) ist ein gutes Beispiel für einen solchen Satz.
Es gibt nichtparadoxe Sätze, die denselben Wahrheitswert in jedem Fixpunkt haben, aber trotzdem keinen wesentlichen Wahrheitswert haben. Z.b.

(13)$P \vee \neg P$

wobei P ein ungegründeter, nichtparadoxer Satz ist. Dann ist (13) wahr in einigen Fixpunkten, and falsch in keinem. Angenommen es gibt Fixpunkte, die P wahr machen, sowie andere die P falsch machen.(z.b. P ist (3)) Dann kann (13) keinen wesentlichen Wahrheitswert haben.
Es gibt keinen gröšten" Fixpunkt, der alle anderen übersteigt. Je zwei Fixpunkte, die demselben Satz verschiedene Wahrheitswerte zuordnen besitzen keine gemeinsame Erweiterung.

[13]Ang. $\mathcal{L}_0 = \mathcal{L}((9), \emptyset)$, dann ist $\mathcal{L}_1 = \mathcal{L}(S, T)$, wobei $S = \{T((9)), \neg T((9)), ...\}$ ist. Widerspruch. Wenn (9) als falsch initialisiert wird, d.h. $\mathcal{L}_0 = \mathcal{L}(\emptyset, (9))$, dann ist $\mathcal{L}_1 = \mathcal{L}(S, T)$, wobei $S = \{T((9)), \neg T((9)), ...\}$, $T = \{T((9)), ...\}$, weil $T((9))$ Element von T ist, $\neg T((9))$ Element von S ist, und somit $T(\neg(9))$ Element von S ist, woraus folgt, dass $(\exists x)(P(x) \wedge T(x))$ Element von S ist, da aber (9) der einzige Satz mit der Eigenschaft $P(x)$ ist, folgt, dass $T((9))$ Element von S ist.

[14]Ein maximaler Fixpunkt ist eine Menge von Sätzen, die inkonsistent wird wenn man einen Satz hinzufügt der noch nicht in der Menge enthalten ist.

Doch es gibt eine größten wesentlichen Fixpunkt (die wesentlichen Fixpunkte bilden einen Verband unter $\leq$). Der größte wesentliche Fixpunkt ist die eindeutige "größte" Interpretaion von $T(x)$, die keine unwesentlichen Wahrheitswerte vergibt.

Um Kripkes Modell mit dem Tarskis vergleichen zu können, muss Tarskis Hierachie auf transfinite Ebenen erweitert werden.
Da diese Erweiterung recht komplex ist, zieht Kripke seinen Vergleich nur für endliche Ebenen.
Sei $A_1(x)$ ein syntaktisches Prädikat, dass genau auf die Formeln zutrifft, die $T(x)$ nicht enthalten (d.h. auf alle Formeln von L). Definiere $T_1(x)$ als $T(x) \wedge A_1(x)$. Sei $A_2(x)$ ein Prädikat, dass genau auf die Formeln zutrifft, die $T_1(x)$ enthalten oder in L sind. Interpretiere $T(x)$ als kleinsten Fixpunkt $\mathcal{L}_\sigma$, dann können alle (totalen) Prädikate der endlichen Tarski-Hierachie in $\mathcal{L}_\sigma$ definiert werden, und alle Sprachen dieser Hierachie sind Subsprachen von $\mathcal{L}_\sigma$.

Was Kripke als Modallogiker auch interessiert, ist wie modale Operatoren in diese Theorie integriert werden können. Dazu muss das Wahrheitsprädikat$T(x)$ in jeder möglichen Welt bewertet werden.
Eine andere Anwendung der Theorie ist die "nichtpredikative" substitutionelle Quantifikation, z.b. in einer Sprache mit substitutionellen Quantoren, die beliebige Sätze der Sprache als Substituenden enthält. In klassische Sprachen können solche Quantoren nicht ohne Wahrheitswertlücken eingeführt werden.
Kripke merkt an, dass ironischerweise seine Theorie für jene von Interesse ist, die intensionale/modale Operatoren und mögliche Welten nicht lieben, und die Modalitäten und propositionale Einstellungen als Prädikate von Sätzen oder Satztoken verstehen.
Wenn ein Notwendikeitsoperator und ein Wahrheitsprädikat erlaubt sind, dann können wir ein Notwendigkeitsprädikat entweder durch $T(nec(x))$ oder durch $\Box T(x)$ definieren, wobei die erste Version Vertretern der Tokentheorie besser gefallen wird wie K. anmerkt.

Ausserdem erwähnt Kripke so nebenbei, dass Montague und Kaplan gezeigt haben, dass dieses Verständnis von modalen Operatoren zu Paradoxien führt, ähnlich der Lügnerparadoxie. Ich möchte hier kurz auf dieses Argument eingehen, und beziehe mich dabei auf das Buch von Vann McGee [McGee:1991] der sich wiederum auf Richard Montague [Montague:1974] bezieht.

Montague hat folgendes Theorem bewiesen.

Theorem 4.2.1 *Sei* Γ *eine Theorie für die gilt:*

(1) Γ *enthält eine Theorie der eigenen Syntax.*
(2) Gamma ist unter logischer Folgerung abgeschlossen.
(3) Γ *enthält* $\phi(\ulcorner\chi\urcorner)$ *wenn* Γ ϕ *enthält.*
(4) Γ *enthält alle Instanzen des Schemas* $\phi(\ulcorner\chi\urcorner) \to \chi$.

Dann ist Γ *inkonsistent.*

Beweis 4.2.1 *Sei* λ *eine Formalisierung des Lügnerparadoxons.*

$(i)\neg\phi(\ulcorner\lambda\urcorner) \to \lambda$
(i) *ist ein Theorem von* Γ *aufgrund der ersten Bedingung.*
$(ii)\phi(\ulcorner\lambda\urcorner) \to \lambda$
(ii) *ist eine Instanz von (4).*
$(iii)\lambda$
$(iv)\phi(\ulcorner\lambda\urcorner)$
(iv) *folgt aus* (iii) *mittels (3).*
$(v)\phi(\ulcorner\lambda\urcorner) \to \neg\lambda$
(v) *ist ein Theorem von* Γ *aufgrund der ersten Bedingung.*
$(vi)\neg\lambda$

(i) ist eine Instanz des Diagonallemmas. Für jedes Prädikat, also auch das Prädikat "nicht notwendigerweise wahr zu sein" gibt es einen Satz der von sich selbst behauptet dieses Prädikat zu erfüllen. (v) folgt aus (i) mittels Theoremen bezüglich des Bikonditionals.
Die ersten beiden Bedingungen sollten von jeder interssanten Theorie erfüllt werden, und werden tatsächlich auch von konsistenten Theorien erfüllt. Die zweite Bedingung ist unerlässlich, die erste Bedingung wird schon von einfachsten mathematischen Theorien erfüllt. Sie werden also nicht für den Widerspruch verantwortlich gemacht werden.

Die dritte und vierte Bedingung sollten von einem Notwendigkeitsprädikat erfüllt werden. Wenn ein Satz ableitbar ist, dann sollte er auch notwendig sein, da es sich um eine Theorie der logischen Notwendigkeit handelt. Diese Bedingung entspricht der *rule of necessitation* oder *Regel zur Einführung des Notwendigkeitsoperators* in der Modallogik. Und wenn ein Satz notwendigerweise wahr ist, dann sollte er auch wahr sein. Dieses Axiomenschemata wird

von allen Modallogiken verwendet.
Diese beiden Bedingungen werden für den Widerspruch verantwortlich gemacht werden, woraus folgt dass es keine Theorie der logischen Notwendigkeit, in der Notwendigkeit als Prädikat von Sätzen verstanden wird, geben kann.
In der Modallogik besteht der Ausweg darin "Notwendigkeit" als Operator zu verstehen und nicht als Eigenschaft von Sätzen. Das legt den Schluss nahe "Wahrheit" auch als Operator und nicht als Eigenschaft von Sätzen zu verstehen. Wie aber McGee[15] gezeigt hat, führt das zu einer Definition die ω-inkonsistent ist.

Es gibt einen Schwachpunkt von Kripkes Theorie, der von ihm selbst kurz erwähnt wird und auch von Tyler Burge kritisiert wird. K. meint, die Arbeit an Sprachen mit Wahrheitsw ertlücken war vielleicht auch motiviert von der Suche nach einer universalen Sprache, einer Sprache in der alles was behauptet werden kann auch ausgedrückt werden kann. Die Theoreme von Gödel und Tarski gelten nur für "klassische" Sprachen, widerlegen dieses Ansatz also nicht prinzipiell.

Die hier untersuchten Sprachen enthalten zwar ein Wahrheitsprädikat und ein Erfüllbarkeitsprädikat, die Induktion welche den Fixpunkt definiert ist allerdings in einer mengentheoretischen Metasprache durchgeführt worden.

In der Metasprache können Aussagen über die Objektsprache gemacht werden, die nicht in der Objektsprache selbst ausgedrückt werden können. Lügnersätze sind nicht wahr in der Objektsprache in dem Sinn, dass sie durch den induktiven Prozess nie wahr gemacht werden. Aber in der Objektsprache können wir das nicht ausdrücken, wegen unserer Definition der Wahrheit und der Negation.

Die Falschheit des Lügnersatzes kann nur behauptet werden von einer Position jenseits des induktiven Prozesses. K. meint lapidar, dass eine universale Sprache wahrscheinlich nicht zu erreichen sei, hält aber die Notwendigkeit auf eine Metaebene aufzusteigen für eine der Schwächen seiner Theorie. "The ghost of the Tarski hierachy is still with us"[16]

Die Metasprache die Kripke verwendet hat keine Wahrheitswertlücken. Entweder ein Satz hat einen Wahrheitswert an einem Fixpunkt oder er hat kei-

[15]Siehe [McGee:1991], Seite 29.
[16]Siehe [Kripke:1984], Seite 80

nen. Begriffe wie "gegründet" und "paradox" gehören zur Metasprache.

K. hält diese Situation für intuitiv gerechtfertigt dadurch, dass sich die obigen Begriffe nicht in der natürlichen Sprache finden solange bis Philosophen über ihre Semantik nachdenken. Er hält sein Modell für ein plausibles Modell der natürlichen Sprache, vorausgesetzt die Idee einer universellen Sprache wird aufgegeben. Ausserdem glaubt er, dass die Konstruktion der Fixpunkte den Lernprozess reflektiert, also die Verwendung des Wahrheitsprädikats im alltäglichen Leben des nichtphilosophischen Sprechers.

Obwohl es ihm nicht gelingt dem Geist der Tarski-Hierachie zu entkommen, hält K. seinen Zugang für intuitiv was die Anwendung auf die natürliche Sprache betrifft. Und wenn man die erweiterten Möglichkeiten der Analyse bedenkt, stimmt das auch. Die Frage ist nur, wie relevant die analysierten Sätze in der natürlichen Sprache tatsächlich sind.

Kripke hat in seiner Theorie folgende Version von Tarskis "Konvention T" verwendet. Sei 'k' der Name eines Satzes A, $T(k)$ ist wahr bzw. falsch genaudann wenn A wahr oder falsch ist. $T(k)$ hat keinen Wahrheitswert wenn A keinen hat.

Eine ander Möglichkeit wäre, wenn A falsch oder undefiniert ist, $T(k)$ als falsch zu bewerten und seine Negation als wahr. An einem beliebigen Fixpunkt werden dann alle nichtwahren Sätze als falsch bewertet. $T(x)$ wird abgeschlossen. In diesem Fall gilt eine modifizierte Version von Tarskis Konvention T.

$$T(k) \vee T(neg(k)). \rightarrow .A \leftrightarrow T(k)$$

Wenn A ein paradoxer Satz ist, dann kann $\neg T(k)$ behauptet werden.

Abschliessend möchte ich Kripkes Arbeit noch in einen grösseren Kontext stellen und bewerten. Mit der Idee dass "Wahrheit" ein riskanter Begriff ist, dass wir also nicht wissen können ob eine paradoxe Situation vorliegt oder nicht, hat Kripke eine spezielle Art von Skeptizismus formuliert bzw. angesprochen. In seiner radikalsten Variante könnte man diesen Skeptizismus so formulieren.

> Wenn ich einen Satz behaupte, kann ich nie sicher sein dass dieser Satz überhaupt konsistent bewertet werden kann

Einen Ausweg aus dieser Situation soll der Begriff der "Gegründetheit" liefern. Dabei taucht allerdings das Problem auf, dass prinzipiell alle empirischen Fakten in die Bewertung eines Satzes einbezogen werden müssen. Alle empirischen Fakten sind für die Bewertung eines Satzes relevant. Unabhängig davon ob das überhaupt möglich ist werden die Modelle in denen ein Satz bewertet werden kann damit extrem eingeschränkt. Jede Änderung in den empirischen Fakten ist eine Änderung des Modells, was eine Neubewertung des Satzes verlangt. Das skeptische Problem ist damit nur verschoben worden.

K. kann zwar den von ihm formulierten Einwänden begegnen, schafft damit aber ein Problem, dass innerhalb seiner Theorie nicht gelöst wird. Wie kann eine Lösung dieses Problems aussehen? In jedem Fall müssen die empirischen Fakten die zur Bewertung eines Satzes relevant sind, eingeschränkt werden. Dafür benötigt man den Begriff einer Sprechergemeinschaft oder einen funktionell ähnlichen Begriff. K.' s Theorie kann zwar modifiziert werden, sodass sie auch einen solchen Begriff enthält, im Prinzip besteht aber die Möglichkeit ohne eine solche Einschränkung auszukommen, was zu Problemen führt. Eine andere Möglichkeit wäre, die Unterscheidung zwischen empirischen und anderen Fakten bzw. Sätzen überhaupt fallenzulassen.

Warum ist diese Theorie aber trotzdem interessant? Formal gesehen ist K.'s Theorie interessant, weil sie es ermöglicht für alle Sprachen die ihre eigene Syntax kodieren können ein Wahrheitsprädikat zu definieren dass in dieser Sprache liegt.

Bezüglich einer Analyse der natürlichen Sprache ist sie von Interesse, weil mehr Sätze analysiert werden können als mit Tarski's Theorie.

Kapitel 5

Tyler Burge

5.1 Die Einwände gegen Kripke

Der Ansatz von Tyler Burge [Burge:1984] ergibt sich aus einer Kritik an Kripkes Theorie und an Theorien die Wahrheitswertlücken verwenden allgemein.
Auch Kripke arbeitet mit Wahrheitswertlücken, da er ja mehrmals betont keinen dritten Wahrheitswert verwendet zu haben. Ein Satz hat nicht den Wahrheitswert "unbestimmt" sondern er hat keinen Wahrheitswert. Vom Wahrheitsprädikat aus betrachtet existiert an dieser Stelle eine Lücke.

Kripkes Kritikpunkt an Tarski war, dass bestimmte Sätze die intuitiv verstanden und bewertet werden können in der Theorie Tarskis nicht bewertet werden können.
Tyler Burge bringt zwei andere Kritikpunkte vor.
Tarski kommt zu dem Ergebnis, dass der Wahrheitsbegriff der natürlichen Sprache inkonsistent ist, und durch einen Wahrheitsbegriff ersetzt werden muss dessen Postulate Tarski angibt. Daraus folgt, dass die natürliche Sprache inkonsistent ist. Um das zu beweisen müssen allerdings Postulate identifiziert werden, die diesem Begriff zugrunde liegen. Das aber hält Burge für illegitim, da Postulate Teile von Theorien sind die von Menschen vertreten werden, und nicht Teile von natürlichen Sprachen *an sich*.
Der andere Kritikpunkt an Tarski ist in Bezug auf Tarski derselbe den auch

Kripke vorbringt, in Bezug auf Kripke[1] meint Burge aber er hätte es nicht geschafft eine intuitive Theorie der Wahrheit zu formulieren, sondern er hätte sich hauptsächlich auf die technische Raffinesse der Theorie konzentriert.

Allen Theoretikern die Wahrheitswertlücken verwenden um mit den semantischen Antinomien und anderen Problemen zu Rande zu kommen wirft B. vor ähnliche Probleme auf einer anderen Ebene zu erzeugen. Die Probleme werden lediglich verschoben. Das bekannteste dieser Probleme ist "der hartnäckige Lügner".[2] Der hartnäckige Lügner lautet in seiner einfachsten Form:

(1)(1) ist nicht wahr.

Wenn wir diesen Satz als weder wahr noch falsch charakterisieren, dann folgt daraus dass dieser Satz *nicht wahr* ist. Also ist (1) nicht wahr. Damit haben wir aber aber einen Satz behauptet, von dem wir angenommen haben er sein weder wahr noch falsch.
K. spricht dieses Problem in seinem Aufsatz beiläufig an. Er sagt, der Lügner sei nicht wahr in seiner Objektsprache und dieser Sachverhalt kann nicht in der Objektsprache ausgedrückt werden. Der Sachverhalt kann in einer zweiwertigen Metasprache ausgedrückt werden. Wenn es aber nur einen Wahrheitsbegriff der natürlichen Sprache gibt, dann kann K.s Analyse nicht korrekt sein. Da K. sehr wenig zu diesem Problem sagt wendet sich B. anderen Autoren zu. Er unterscheidet zwischen drei Typen von möglichen Lösungen im Kontext einer Wahrheitswertlückentheorie.[3] Er versucht zu zeigen, dass keine der drei eine Lösung des Problems ist.
Man könne nicht behaupten dass (1) nicht wahr ist, wenn man die Rede von Wahrheitswertlücken ernst nimmt. [Mackie:1973] Die Metasprache enthält dieselben Wahrheitswertlücken wie die Objektsprache. B. hält diesen Ansatz zwar für sehr konsequent aber nicht glaubhaft. Ziel dieses Ansatzes ist, die Folgerung von "(1) ist weder wahr noch falsch" zu "(1) ist nicht wahr" abzublocken. B. hält diese Folgerung aber für intuitiv und nicht angreifbar.[4]

[1]Burge bezieht sich nicht direkt auf Kripke sondern auf "post-tarskian treatments of the paradoxes" allgemein, womit wohl auch K. gemeint ist.

[2]"The persistent Liar" oder "The strengthened Liar".

[3]"truth-value gap theory"

[4]Diese Folgerung kann wahrscheinlich nur verhindert werden, wenn "wahr" in einem erkenntnistheoretischen Sinn verstanden wwird.

Darüber hinaus gibt es aber auch ein anderes Problem. Der hartnäckige Lügner wird zum hartnäckigsten Lügner.

(2) (2) ist entweder falsch oder undefiniert.

Aus (2) folgt, dass (2) nicht wahr ist. Das ist aber gerade dass was (2) behauptet.
Eine andere "Lösung" wäre, die Substituitivität der Identität einzuschränken und so die Folgerung von "Dieser Satz ist falsch" zu "(1) ist falsch" zu unterbinden. [Skyrms:1970] Diese Lösung hält B. für *ad hoc* und ungenügend da auch Paradoxien formuliert werden können ohne die Substitution zu verwenden.
Die dritte von B. erwähnte Lšung besteht darin, den Begriff der Negation genauer zu untersuchen. Diese Variante greift ebenfalls zu kurz, denn es gibt Paradoxien die ohne die Verwendung der Negation formuliert werden können.Da die Verwendung von Wahrheitswertlücken keine Lösung der Probleme bietet schlägt B. vor das Konzept der Hierachie von Prädikaten neu zu überdenken. Er übernimmt allerdings von den "truth-value gap theorists" die Idee "ist falsch" nicht mit "ist ein Satz und ist wahr" zu identifizieren.

B. will die Folgerungen, die zum "hartnäckigen Lügner" führen interpretieren um sie zu rechtfertigen. Dabei geht es um (a) das Auftauchen des "Lügners", woraus folgt, dass dieser Satz pathologisch ist was zur Konklusion (b) führt, dass dieser Satz nicht wahr ist. Daraus folgt (c), d.h. der Satz ist trotz allem wahr. Um diese Folgerungen zu untersuchen verwendet B. ein Beispiel, dass er von A.N. Prior [Prior:1961] übernommen hat.

> (C) Angenommen ein Student der glaubt, dass er sich im Hörsaal 10 befindet und dass der Vortragende in Hörsaal 9 nicht sehr gut ist, schreibt den Satz (a)"Es gibt keinen wahren Satz auf der Tafel im Hörsaal 9 am 12.12.1999" auf die Tafel. Unglücklicherweise ist der 12.12.1999, der Student befindet sich im Hörsaal 9 und der einzige Satz der dort auf der Tafel steht ist (a). Daraus folgern wir, dass der Satz nicht wahr ist. Das heisst, dass (b) es keinen wahren Satz auf der Tafel in Hörsaal 9 am 12.12.1999 gibt. Damit haben wir aber gerade den fraglichen Satz behauptet und folgern (c) dass er wahr ist.

Burge verwendet für seine Analyse den Begriff der Implikatur. Eine Implikatur gibt für ein semantisches Prädikat an, wie es in einem bestimmten

Kontext verwendet werden soll. Es handelt sich also um einen Begriff der Pragmatik. Auf das obere Beispiel angewendet entscheidet die Implikatur welches Wahrheitsprädikat verwendet werden soll.[5] Die relevante Implikatur für das Wahrheitsprädikat lautet:

> *Sätze auf die referiert oder über die quantifiziert wird, müssen mit dem Wahrheitsprädikat des bewertenden Satzes bewertet werden.*

Dadurch dass in dem Beispiel der bewertende und der bewertete Satz derselbe sind, entsteht bei Anwendung der Implikatur ein Paradox.

5.2 Die Lösung

5.2.1 Analyse der Folgerung von (a) auf (b)

Burge möchte zeigen, dass in der Folgerung von (a) auf (b) keine Änderung des Wahrheitswerts oder der semantischen Bewertung involviert ist, sondern dass sich lediglich die Implikatur bzw. die Hintergrundannahmen des Sprechers bzw. Interpreten ändern.
B. hält seine Ansicht mit der Auffassung für kompatibel, die meint paradoxe Sätze schaffen es nicht eine Proposition auszudrücken, und hätten daher keinen Wahrheitswert. Das ein Satz daran scheitert eine Proposition auszudrücken bedeutet dann, dass vom Wahrheitsprädikat welches bezüglich der Implikatur passend ist keine Wahrheitsbedingungen angegeben werden. B. sieht darin eine Annäherung zwischen seiner Auffassung und den "truth-value-gap"-Theorien. Einige Sätze sind nicht wahr, nicht weil ihre "$wahr_i$"-Wahrheitsbedingungen nicht erfüllt sind, sondern weil sie keine solchen Wahrheitsbedingungen haben. Der Unterschied zu den "truth-value-gap-Theoretikern besteht für B. darin, dass in seiner Theorie dieses Fehlen von Wahrheitsbedingungen nichts Absolutes ist.

Behauptet man, dass ein paradoxer Satz es nicht schafft eine Proposition auszudrücken, dann muss man zwei Einschränkungen beachten.

1. Proposition kann nicht im Sinne von "etwas dass geglaubt werden kann" oder "etwas dass Bedeutung oder Sinn hat" verstanden werden.

[5] B. meint, dass in dem Beispiel ein indexikalisches Element am Werk ist, und das Element dem man die Indexikalität zuschreiben muss das Wahrheitsprädikat ist.

2. Diese Auffassung kann nicht erklären warum sich der Wahrheitswert eines Satzes wie im Beispiel (C) von "drückt keine Proposition aus" zu "wahr" verändert.

Wie aber kann diese Veränderung verstanden werden?
Versucht man sie als Veränderung der Extension des Wahrheitsprädikats zu verstehen gerät man in Probleme. Angenommen wir machen das verwendetete Wahrheitsprädikat von (a) aus Beispiel (C) explizit, formulieren also (a) um zu (α)"Kein Satz der auf der Tafel...ist $wahr_i$." Der Übergang zu (b) würde eine Veränderung der Extension beinhalten: (β)"Kein Satz der auf der Tafel...ist $wahr_k$." Aber wir wollen ja wissen ob (α) $wahr_i$ ist oder nicht wahr. Beantworten wir diese Frage mit "ja" erhalten wir einen Widerspruch, beantworten wir sie mit "nein" dann können wir (b) als (β)"Kein Satz der auf der Tafel...ist $wahr_i$." reformulieren. Damit kann es aber keine Veränderung im Wahrheitswert zwischen (α) und (β) geben, weil sie ein und derselbe Satz sind.
Entweder gibt es keine Veränderung in der semantischen Bewertung zwischen (a) und (b) oder "wahr" ist nicht extensional. Da es darüber hinaus keinen Grund gibt "wahr" als nicht extensional aufzufassen hält B. diese Lösung für *ad hoc.*

Diese Überlegungen führen Burge zu der Auffassung, dass die Veränderung von (a) zu (b) in pragmatischen Begriffen verstanden werden muss und nicht in semantischen Begriffen. Was sich bei der Bewertung verändert sind unsere Hintergrundannahmen über das Wahrheitsprädikat das verwendet werden muss. Das Paradox entsteht sozusagen aus falschen Erwartungen.
Diese Herangehensweise erklärt warum es so scheint als würde sich der Wahrheitswert des Satzes verändern. (a) ist pathologisch und nicht $wahr_i$, (b) ist $wahr_k$, aber der Satz ist an beiden Stellen nicht $wahr_i$ und $wahr_k$.

5.2.2 Analyse der Folgerung von (b) auf (c)

Die Folgerung von (b) auf c) analysiert B. anhand der Indexikalität von sematischen Prädikaten. Diese Indexikalität wird auf zwei verschiedene Arten analysiert. Die Verwendung der Indices wird durch die strukturelle Analyse erläutert, wie in einem Kontext der passende Index gefunden werden kann gehört zur materiellen Analyse.

5.2.2.1 Strukturelle Analyse der Indexikalität

Die Extension der indexikalischen Prädikate verändert sich je nach Kontext. Sie werden aber nicht als Variablen verstanden, da nicht über diese Prädikate quantifiziert wird. "Wahr" wird als schematisches Prädikat verstanden, dass in einem gegebenen Kontext eine bestimmte Extension annimmt und in diesem Kontext wird "wahr" durch ein Prädikat "$wahr_i$" mit einem numerischen Index repräsentiert.
B. meint die Hauptidee der indexikalischen Analyse sei die Definition des Begriffes eines pathologischen Satzes, d.h. die Definition eines Prädikats "$pathologisch_i$". Da mir "pathologisch" in diesem Zusammenhang zu dramatisch klingt und ausserdem auf eine wesentlich problematische Eigenschaft eines Satzes hindeutet, etwas das Burge ja gerade widerlegen will, werde ich das Wort "seltsam" verwenden. Sätze die $seltsam_i$ sind, sind möglicherweise nicht $seltsam_k$ für $k > i$. Seltsamkeit ist keine wesentliche Eigenschaft eines Satzes sondern eine Disposition bei bestimmten Bewertungen Seltsames zu produzieren.
Um das Beispiel (C) strukturell zu analysieren entwickelt B. drei Konstruktionen (C1)-(C3). Um Instanzen der Axiome dieser Konstruktionen zu interpretieren muss die Interpretation den nächsthöheren Index des Wahrheitsprädikats verwenden. Die Basisidee aller drei Konstruktionen ist "to define a notion of a $pathological_i$ sentence."[6]

Konstruktion (C1)

Die erste Konstruktion ist analog der Konstruktion Tarski's. Alle Sätze sind $seltsam_i$, die ein Wahrheitsprädikat enthalten, dessen Index grösser ist als i. Diese Sätze sind nicht $wahr_i$. Nur Sätze die nicht $pathologisch_i$ sind, können Instanzen eines $wahr_i$-Schemas sein. B.'s Konstruktion unterscheidet sich von derjenigen Tarski's darin, dass die Anwendung eines $wahr_i$-Prädikats auf Sätze die $seltsam_k$, $k >= i$ sind wohlgeformt ist, was durch eine Konditionalisierung des Wahrheitsschemas erreicht wird. Damit können Prädikationen untersucht werden, die von Tarski schon auf der syntaktischen Ebene ausgeschlossen werden. B. hält aber diese erste Bedingung für eine zu starke Einschränkung. Um die Klasse der bewertbaren Sätze zu erweitern entwickelt B. die zweite Konstruktion.

[6] [Burge:1984], Seite 101.

Konstruktion(C2)

$$(1) S_i(\ulcorner Sat_k(t,t_1)\urcorner, \alpha), S_i(\ulcorner S_j(t,t_1)\urcorner, \alpha) k, j >= i$$

"Sat_k" ist wieder das Erfüllbarkeitsprädikat, α ist eine Bewertung. "$seltsam_i$" (d.h. S_i) wird relativ zu einer bestimmten Bewertung α definiert.[7] B. relativiert die Definition, weil die Seltsamkeit mancher Sätze von der Bewertung abhängt. (1) bedeutet, dass alle Sätze $seltsam_i$ sind, die ein Erfüllbarkeitsprädikat oder ein Seltsamkeitsprädikat mit höherem Index als i enthalten, die formale Variante von Konstruktion (C1).

$$(2)\ if\ S_i(\Gamma, \alpha),\ then\ S_i(\ulcorner \neg\Gamma \urcorner, \alpha)$$

Wenn ein Satz seltsam ist, dann ist auch seine Negation seltsam.

Die nächsten beiden Punkte beziehen sich auf die Seltsamkeit von Sätzen in denen Quantoren oder ausagenlogische Verknüpfungen vorkommen. Diese beiden Punkte werde ich nicht in der formalen Variante wiedergeben, weil ich glaube dass es einfacher ist ihren Zweck zu erläutern.

(3) Konditionale werden nur dann seltsam, wenn ein Teil des Konditionals seltsam ist, und das Konditional auf der entsprechenden (indexikalischen) Ebene nicht erfüllbar ist. Ein Konditional dessen Vorderglied falsch ist, dass somit $erfuellbar_i$ ist, ist nicht $seltsam_i$, unabhängig von den Eigenschaften der restlichen Teile des Konditionals. Der Begriff der Seltsamkeit ist wesentlich schwächer als der Begriff der Erfüllbarkeit oder der Begriff der Wahrheit. Ist ein Satz einmal wahr, dann ist er sozusagen immunisiert gegen Seltsamkeiten.

(4) Wenn es eine Bewertung α gibt, unter der ein Satz Φ $seltsam_i$ ist, und es keine Bewertung α_1 gibt sodass $\alpha =^x \alpha_1$ und α_1 Φ falsifiziert, dann ist $\forall x \Phi$ $seltsam_i$ bezüglich α. Angenommen wir haben irgendein Prädikat. Wenn das Prädikat nicht auf alles zutrifft, dann ist die Allquantifikation des Prädikats falsch. Die Allquantifikation kann nur dann $seltsam_i$ werden, wenn das Prädikat bezüglich einigen Gegenständen seltsam wird, und auf alle anderen zutrifft. Was den Allquantor betrifft, ist also der Begriff der Falschheit wesentlich stärker als der Begriff der Seltsamkeit.

Wie bei Punkt (1) und (2) wurden auch (3) und (4) relativ zu einer Bewertung definiert. Der fünfte Punkt ist die induktive Abschlussregel für das Prädikat $seltsam_i$.

---

[7] Eine Bewertung ist eine Funktion, die allen Variablen Werte zuordnet. Relativ zu einer Bewertung α können dann auch Formeln mit freien Variablen ausgewertet werden. $\alpha =^x \alpha_1$ bedeutet, dass die beiden Bewertungen bis auf den Wert von x übereinstimmen.

In Punkt (6) wird eine Verbindung hergestellt zwischen den Prädikaten $seltsam_i$ und $erfuellbar_i$.

$$(6) S_i(\Gamma, \alpha) \rightarrow \neg Sat_i(\Gamma, \alpha)$$

Wenn ein Satz seltsam ist, dann ist er nicht erfüllbar. Ein interessanter Punkt an B.'s Theorie ist, dass er "Seltsamkeit" und "Erfüllbarkeit" als durcheinander bestimmte Begriffe ansieht. Unter Verwendung des Begriffs der Seltsamkeit werden hinreichende Bedingungen für Erfüllbarkeit (oder Nicht-Erfüllbarkeit) angegeben (Punkt (6)) und umgekehrt (Punkt (3) und (4)). Beide Begriffe werden als grundlegend und zusammengehörig angesehen.

In den nächsten drei Punkten (7), (8) und (9), die nach meiner Auffassung das Kernstück der ganzen formalen Konstruktion ausmachen wird eine rekursive Charakterisierung der "Wahrheit" auf Basis der Nicht-Seltsamkeit gegeben. Es handelt sich dabei um keine Definition von "Wahrheit".

(7) Wenn die Negation eines Satzes nicht $seltsam_i$ ist, dann ist die Negation erfüllbar genau dann wenn der Satz nicht erfüllbar ist.

(8) Wenn ein Konditional nicht $seltsam_i$ ist, dann ist das Konditional erfüllbar genau dann wenn unter der Bedingung dass das Antecedens erfüllbar ist auch das Konsequens erfüllbar ist.

(9) gibt die Regel für den Allsatz an, d.h. sie entspricht genau der klassichen Regel ist lediglich auf Nicht-Seltsamkeit relativiert.

Es lässt sich auch noch ein eingeschränktes Wahrheitsschema definieren:

$$(T) \neg S_i(\beta, \alpha) \rightarrow (Sat_i(\beta, \alpha) \leftrightarrow P)$$

Was ist damit erreicht?

Eine rekursive Charakterisierung von "Wahrheit" und ein eingeschränktes Wahrheitsschema mit dem der Lügner nicht bewertet werden kann.[8] Da die Begriffe "Seltsamkeit" und "Erfüllbarkeit" durcheinander definiert sind, ist die Definition allerdings zirkulär. Um "Wahrheit " zu bestimmen, müssen wir nicht nur "Erfüllbarkeit" voraussetzen sondern auch "Seltsamkeit", d.h. ein Verständnis davon ob ein Satz überhaupt bewertet werden kann. Dieses Verständnis kann sich aber nicht nur auf syntaktische Eigenschaften des Satzes beziehen, sondern muss auch semantische Eigenschaften umfassen.

Wie charakterisiert Burge seine Konstruktion?

[8] Unter bestimmten Bedingungen, siehe weiter unten.

Ein positives Ergebnis seiner Theorie sieht Burge darin, dass für alle drei Konstruktionen[9] das Prinzip des ausgeschlossenen Dritten $\ulcorner P \vee \neg P \urcorner$ gilt. Alle geschlossenen Sätze sind entweder $wahr_i$ oder nicht $wahr_i$ für alle i. Es gilt jedoch nicht, dass entweder ein Satz oder seine Negation $wahr_i$ sein muss.

Durch die Indexikalität des Wahrheitsprädikats kann auch dem Vorwurf der Mehrdeutigkeit, den Burge gegen Tarski vorgebracht hat begegnet werden. Dazu Burge:

> *In natural language there is a single indexical predicate. We represent this predicate by the schematic predicate expression* $\ulcorner true_i \urcorner$. [Burge:1984, S.107]

Durch die Kontextualisierung der Indices entstehen keine neuen Wahrheitsprädikate sondern *contextual applications of the indexical 'true'* [Burge:1984, S.107]

Der Index darf aber in einer semantischen Untersuchung[10] nicht als eine freie Variable verstanden werden, über die quantifiziert werden kann. Dann könnte man nämlich einen "Super Liar" [Burge:1984, S.108] konstruieren, einen Satz der von sich behauptet auf keiner Ebene wahr zu sein. Burge meint, dass auch dieser Satz einen versteckten Index hat.

5.2.2.2 Materielle Analyse der Indexikalität

Die materielle Analyse soll die Basis liefern für die Anwendung des indexikalischen Wahrheitsprädikats. Dabei geht es um die Frage:

> *The chief question about the application of the formal structure is how the subscripts are established in context.* [Burge:1984, S.108]

Welcher Index des Wahrheitsprädikats soll bei einer Anwendung auf die natürliche Sprache verwendet werden?

---

[9] Konstruktion 3 habe ich nicht besprochen, weil sie nur eine Verfeinerung von Konstruktion 2 ist.

[10] Nicht so bei anderen Untesuchungen. Dazu Burge:

> *When we are given a semantical theory for nonsemantical indexical sentences, we relativize the semantical predicate to the context so as to generalize over all possible uses of the relevant indexical expressions* [Burge:1984, S.108]

In einer ersten Annäherung an diese Frage plädiert Burge für ein "Principle of Verity", welches die Fähigkeit des Interpreten einem Satz einen Wahrheitswert zu geben maximieren soll.[11]

Zur pragmatischen Rechtfertigung dieses Prinzips meint Burge: "it excuses us in ordinary discourse from worrying about paradox,.." [Burge:1984, S.110] Die Extension des Wahrheitsprädikats hängt damit auch vom Kontext der Verwendung ab, was Das zweite Prinzip der materialen Analyse ist das "Principle of Justice",das Burge folgendermassen formuliert:

> *One should not give one statement truth conditions instead of another without some reason.* [Burge:1984, S.110]

Die Anwendung der beiden Prinzipien zeigt B. anhand deselben Beispiels, das auch schon Kripke verwendet hat, und das ich hier in seiner Originalversion wiedergebe.

> *(A) All Nixon's utterances about Watergate are untrue.*

und

> *(B) Everything Dean utters about Watergate is untrue.*

wobei *(A)* von Dean und *(B)* von Nixon behauptet wird. Da jeder von beiden alle Behauptungen des anderen einbeziehen will sollte nach dem "Principle of Justice" beiden derselbe Index i zugewiesen werden. Nach dem "Principle of Verity" sollte der Index gross genug sein, um alle Aussagen bis auf *(A)* und *(B)* zu umfassen. Hat Dean mindestens eine wahre Behauptung über Watergate gemacht, dann folgt nach Konstruktion 2, dass Nixons Behauptung $nichtseltsam_{i+1}$ ist und nicht $wahr_{i+1}$ ist. Dieselbe Argumentation gilt auch für die Behauptungen Nixon's.

B. unterscheidet zwischen zwei Verwendungsweisen von indexikalischen semantischen Prädikaten. Einer *derivativen* und einer *evaluativen* Verwendung. *Derivative* Verwendung bedeutet, dass semantische Prädikate nur auf Sätze angewendet werden können, die unabhängig von dieser Anwendung einen Sinn und eine Bedeutung haben.

> *As a consequence, no statement can sit in semantical judgement on itself* [Burge:1984, S.113]

[11]Dieses Prinzip muss aber nicht, wie etwa Quine's "Principle of Charity" die Rationalität des Sprechers maximieren.

Sowohl Russel's "visious circle principle" als auch Tarski's Bezug auf eine Metasprache reflektieren diese Verwendungsweise.

Evaluative Verwendung bedeutet, dass bei der Anwendung der semantischen Prädikate die Sätze dahingehend geprüft werden ob "es" sich so verhält wie sie "es" repräsentieren. B. sieht die Tarski'schen Bikonditionale und ihre semantische Analyse als ein Beispiel für diese Verwendung. [Burge:1984, S.114f] Wie Tarski gezeigt hat, ist das Wissen über die Erfüllbarkeitsbedingungen von Prädikaten, Namen usw. eine ausreichende Basis um "wahr" verstehen und verwenden zu können, wenn man aber wissen will ob ein Satz wahr ist, kommt man nicht darum herum die Erfüllbarkeit der Teile des Satzes zu prüfen.

Burge meint, dass in bisherigen Untersuchungen die *derivative* Verwendung des Wahrheitsprädikats eine absolute Begrenzung für die *evaluative* Verwendung war. Solche Ansätze stellen keine Lösungen des "Lügners" dar, weil:

> *..because reflection on the proposed solution ... produces a new evaluation which cannot be expressed in terms of the solution or which is incompatible with it.* [Burge:1984, S.114]

Wenn durch die *derivative* Verwendung die "Basiseigenschaften" der Sätze festgelegt werden, hat man bei der Bewertung des "verstärkten Lügners" Probleme, obwohl man "evaluative Intuitionen" [Burge:1984, S.114] haben kann wie der Satz auszuwerten ist.

Abschliessend möchte ich noch zusammenfassen, was ich für die relevanten Punkte von Burge's Theorie halte.

- Eine Verbindung zwischen den semantischen Begriffen "Seltsamkeit" und "Wahrheit" herzustellen.
- Keine Definiton der Wahrheit zu liefern.
- Den Begriff eines "indexikalischen semantischen Prädikats" ernstzunehmen, d.h. Indices nicht als freie Variablen aufzufassen.
- Eine zirkuläre Bestimmung von semantischen Begriffen zuzulassen.

Kapitel 6

Zusammenfassung

Müsste ich die beiden vorgestellten Theorien anhand eines einfachen Labels unterscheiden, dann würde ich sagen, dass Kripke's Theorie eine "semantische" ist, Burge's Theorie eine "pragmatische". "Definiton oder Nicht-Definition" das markiert den Unterschied. Dieser Unterschied ist allerdings nur das oberflächliche Ereignis, hinter dem andere Entscheidungen stehen.

Das Projekt einer Universalsprache, einer Sprache in der alles was behauptet werden kann auch ausgedrückt werden kann, hält Kripke zwar für gescheitert,[1] aber er ist diesem Projekt doch noch verpflichtet. Burge hingegen ist an diesem Projekt nicht mehr interessiert, sondern begnügt sich damit zu verstehen wie "Wahrheit" funktioniert, semantisch wie pragmatisch. Dabei verwendet Burge den Begriff der *Implikatur*.[2] Es ergibt sich damit ein Zusammenspiel zwischen dem Wahrheitsprädikat und seiner Verwendung. Jemandem, der auf der Suche nach einer Universalsprache ist, wird eine solche Lösung allerdings kaum überzeugen. Durch eine "Semantisierung" der Pragmatik ließe sich zeigen, dass dieser Ansatz wiederum in dieselben Probleme

[1]

> *It seems likely that many who have worked on the truth-gap approach to the semantic paradoxes have hoped for a universal language, ... Now the languages of the present approach contain their own truth predicates ... and thus to this extent the hope has been realized. Nevertheless the present approach certainly does not claim to give a universal language, and I doubt that such a goal can be achieved.* [Kripke:1984, S.79]

[2]Die Art und Weise in der ein sematisches Prädikat verwendet wird, werden soll.

gerät wie ein "semantischer" Ansatz.[3]

Einen Vorschlag in Burge's Richtung macht auch Donald Davidson, wenn er meint man sollte die semantischen Prädikate als "primitiv" (d.h. undefiniert) ansehen.

> *... to honour the recognition that the semantic predicates are primitives, we can drop the final step that for Tarski turns recursive characterisations into explicit definitions, and view the results as axiomatized theories of truth.* [Davidson:1990, S.297]

Dadurch soll eine Beziehung zu den alltagssprachlichen Konzepten von "Erfüllbarkeit" und "Wahrheit" erhalten bleiben. Wird eine Definiton daraus gemacht, dann könnte der Vorwurf kommen, dass alle Einsetzungen ins **T**-Schema, die aus der Definiton folgen Tautologien sind, und daher nichts über eine Sprache sagen.[4]

Burge geht aber noch ein Stück weiter als hier von Davidson gefordert wird. Seine Charakterisierung lässt sich gar nicht in eine explizite Definition umwandeln, weil die Begriffe "Erfüllbarkeit" und "Seltsamkeit" durcheinander definiert sind. Denkt man den Vorschlag Davidson's weiter, und nimmt die Rede von der "Einfachheit" der semantischen Begriffe ernst, dann stellt diese Zirkularität auch kein Problem dar. Es handelt sich nicht mehr um eine problematische zirkuläre Definition, sondern um eine wechselseitige Beziehung zwischen Begriffen.

Sowie durch die Sätze

> *Alle Menschen sind sterblich*

und

> *Nicht jedes sterbliche Wesen ist ein Mensch*

eine wechselseitige Beziehung zwischen den Begriffen "Sterblichkeit" und "Mensch" ausgesagt wird, ohne dass diese Begriffe *durch diese Beziehung*

[3] Nämlich dann, wenn über die Indices des Wahrheitsprädiakts quantifiziert wird.

[4] Hat Tarski nun eine Definition der "Wahrheit" gegeben oder nicht? Er hat gezeigt, dass eine solche Definiton für eine bestimmte Sprache nicht in dieser Sprache selbst gegeben werden kann. In einer Metasprache kann die Defintion aber gemacht werden, und von dieser Definiton ist oben die Rede.

problematisch würden.[5] Ausserdem stehen diese beiden Sätze in keiner einfachen Folgerungsbeziehung, sodass einer der beiden Sätze redundant wäre. Beide zusammen sind wesentlich informativer als einer von ihnen alleine. Eine zirkuläre Beziehung zwischen Begriffen zu analysieren hat durchaus seinen Wert, weil auch Zirkularitäten unterschiedlich interessant sind.

Ein anderer Punkt an dem sich die beiden Autoren unterscheiden ist, wie sie mit der von Tarki übernommenen Differenz zwischen Objekt und Metasprache umgehen. Saul Kripke versucht die Differenz zwischen Objekt und Metasprache zu überwinden indem er das Wahrheitsprädikat der Objektsprache sukzessive durch Referenz auf die Metasprache anreichert, sodass die Objektsprache am Ende dieses idealen weil unendlichen Prozesses ein Wahrheitsprädikat enthält. Betrachtet man diesen unendlichen Prozess allerdings selbst wieder als eine Gesamtheit, als abgeschlossen, dann entstehen ähnliche Probleme wie in Tarski's Theorie.

Für Tyler Burge geht es darum, wie sich eine SprecherIn die ein Wahrheitsprädikat verwendet zwischen diesen verschiedenen Stufen bewegt.

Obwohl Kripke und Burge Kritik an Tarski üben, gehen sie doch in wesentlichen Punkten nicht über seinen Ansatz hinaus. Die wichtigste Idee, die sie von Tarski übernehmen, ist sein Wahrheitsschema. Die Idee, dass eine Klärung dieses Schemas etwas interessantes über "Wahrheit" zutage bringen könnte. In einer ersten Formulierung könnte dieses Schema so gelesen werden

> Die Behauptung eines Satzes ist äquivalent zur Behauptung dass dieser Satz wahr ist.

In dieser Formulierung klingt dass Schema allerdings nicht sehr interessant, allein schon deshalb weil eine Äquivalenz behauptet wird, und vielleicht der Eindruck entsteht dass schon alles klar, weil tautologisch ist. Von daher könnte man vielleicht zu der Ansicht von Richard Rorty kommen,

> *..that truth is not the sort of thing one should expect to have a philosophically interesting theory about ... there is no interesting work to be done in this area.* [Rorty:1982]

[5] Obwohl es problematisch für uns ist sterblich zu sein. Das zeigt unter anderem, dass zwischen einem Begriff und einem Objekt ein Unterschied gemacht werden muss, selbst wenn das Objekt letztlich selbt ein Begriff ist.

Dieser Eindruck verschwindet vielleicht wenn man sieht wie verschieden die beiden Aussagemodi auf beiden Seiten der Äquivalenz sind. Einmal die Behauptung eines Satzes und das andere mal die Erwähnung des Satzes. Der Unterschied zwischen Verwendung und Erwähnung oder zwischen Text und Zitat oder zwischen Bild und Abgebildetem. Und beides zusammengehalten von der Prädikation, womit sichergestellt ist, dass es sich auch um eine Äquivalenz zwischen *Sätzen* handelt, und nicht einfach zwischen Bild und Abgebildetem.

So gesehen ist es schon verwunderlich wie zwischen diesen beiden Modi jemals eine Gleichheit bestehen kann. Diese klassische philospphische Frage finde ich prinzipiell interessant, nicht nur in Bezug auf das Problem der "Wahrheit". Als Tarski's Verdienst kann angesehen werden diese Zusammenhänge auf interessante Weise mit den Mitteln der modernen Logik untersucht und dargestellt zu haben.

Ironischerweise kommt Tarski zu einem ähnlichen Ergebnis wie Rorty bezüglich der philosophischen Bedeutung von "Wahrheit". In seinem Aufsatz *The semantic conception of truth and the foundation of semantics* [Tarski:1944] verteidigt sich Tarski gegenüber Vorwürfen die Semantik arbeite mit metaphysischen Begriffen. Im Rahmen seiner Verteidigung argumentiert Tarski, dass die Semantik zumindest nicht mehr metaphysische Begriffe enthalte als andere Teile der Mathematik oder der theoretischen Physik, und das "metaphysische" Begriffe wie derjenige der irrationalen Zahlen nach einer genauen Analyse ihren metaphysischen Charakter verloren hätten. Und genau diese Analyse habe auch er für den Begriff "Wahrheit" versucht sodass dieser Begriff nach seiner exakten Definition nichts metaphysisches mehr an sich hat.

> *If in consequence semantic concepts loose philosophical interest, they will share the fate of many other concepts of science, and this need give rise to no regret.* [Tarski:1944]

Aber selbst wenn nach der Analyse des Begriffs "Wahrheit" dieser nicht metaphysischer ist als irgendein anderer wissenschaftlicher Begriff, und wir nicht traurig darüber sein müssen, dann gibt es doch etwas philosophisch interessantes von Tarski zu lernen. Und zwar von der Art und Weise wie er den Begriff analysiert und welche Zusammenhänge zu anderen Konzepten hergestellt werden.

Literaturverzeichnis

[Barwise:1977] Jon Barwise: *An Introduction to First-Order Logic*, Handbook of mathematical logic, edited by J.Barwise, North-Holland Publishing, 1977

[Bibel:1985] *Die Bibel*, Deutsche Bibelgesellschaft, Stuttgart, 1985

[Boolos:1990] George Boolos and Richard Jeffrey :*Computability and Logic*, Cambridge University Press, 1990

[Burge:1984] Tyler Burge: *Semantical Paradox*, Recent Essays on Truth and the Liar Paradox, ed. by Robert L. Martin, Oxford University Press, New York, 1984, Seite 83-117

[Cantor:1962] Georg Cantor: *Über eine elementare Frage der Mannigfaltigkeitslehre,Jahresbericht der Deutsch. Math. Vereinig. Bd I, S.75-78, 1890-91*, in Gesammelte Abhandlungen mathematischen und philosophischen Inhalts, Georg Olms Verlagsbuchhandlung, Hildesheim, 1962

[Date:1989] C.J. Date: *Not is not 'Not'*, in Relational Database Writings 1985-1989, Addison-Wesley

[Davidson:1990] Donald Davidson: *The structure and content of truth*, in The journal of philosophy, Volume LXXXVII, No.6, June 1990

[Fitting:1986] Melvin Fitting: *Notes on the Mathematical Aspects of Kripke's Theory of Truth*, Notre Dame Journal of Formal Logic, Volume 27, Number 1, January 1986

[Frege:1966] Gottlob Frege: *Der Gedanke - Eine logische Untersuchung*, Logische Untersuchungen, Vandenhoeck & Ruprecht, Göttingen, 1966

[Goedel:1986] Kurt Gödel: *Über formal unentschiedbare Sätze der* Principia Mathematica *und verwandter Systeme I*, Collected Works, Volume I, Oxford University Press, New York, 1986

[Halbach:1996] Volker Halbach: *Axiomatische Wahrheitstheorien*, Akademie Verlag, Berlin, 1996

[Herzberger:1970] Hans Herzberger: *Paradoxes of Grounding in Semantics*, The Journal of Philosophy, XVII, 6, 1970

[Kaye:1991] Richard Kaye: *Models of Peano Arithmetic*, Clarendon Press, Oxford, 1991

[Kearns:1997] John T. Kearns: *Thinking Machines: Some fundamental Confusions*, Minds and Machines 7, 1997, Seite 269-287

[Kleene:1991] S. C. Kleene: *Introduction to Metamathematics*, Clarendon Press, Oxford, 1991

[Kripke:1984] Saul Kripke: *Outline of a theory of truth*, Recent Essays on Truth an the Liar Paradox, ed. by Robert L. Martin, Oxford University Press, New York, 1984, Seite 53-81

[Kutschera:1985] Franz von Kutschera: *Der Satz vom ausgeschlossssenen Dritten*, Berlin/New York, 1985

[Lemmon:1969] E. G. Lemmon: *Beginning Logic*, Suhrkamp Verlag, Frankfurt am Main, 1969

[Mackie:1973] J.L.Mackie: *Truth, Probability and Paradox*, New York: Oxford, 1973

[McGee:1991] Vann McGee: *Truth, Vagueness, and Paradox*, Hackett Publishing Company, Indianapolis, 1991

[Montague:1974] Richard Montague: *Syntactic Treatments of Modality, with Corollaries on Reflexion Principles and Finite Axiomatizibility*, Acta Philosophica Fennica 16: Seite 153-67, 1974

[Prior:1961] A.N. Prior: *On a Family of Paradoxes*, Notre Dame Journal of Formal Logic, ii, 1 (Jänner 1961)

[Quine:1953] W.V.O. Quine: *Logic and the reification of universals*, From a logical point of view, Harvard University Press, 1953

[Quine:1969] W.V.O. Quine: *Grundzüge der Logik*, Suhrkamp Verlag, Frankfurt am Main, 1969

[Quine:1980] W.V.O. Quine: *Wort und Gegenstand*, Reclam Verlag, Stuttgart, 1980

[Rorty:1982] Richard Rorty: *Consequences of Pragmatism*, Minnesota University Press, 1982

[Shapiro:1990] Stewart Shapiro: *Second-Order Logic, Foundations, and Rules*, The Journal of Philosophy, 1990

[Skyrms:1970] Brian Skyrms: *Return of the Liar* The Paradox of the Liar, ed. by Robert L. Martin, New Haven, 1970

[Smorinski:1977] C.Smorinski: *The Incompleteness Theorems*, Handbook of mathematical logic, edited by J.Barwise, North-Holland Publishing, 1977

[Tarski:1939] Alfred Tarski: *On undecidable statements in enlarged systems of logic and the concept of truth*, Journal of symbolic logic, Volume 4, Number3, September 1939

[Tarski:1944] Alfred Tarski: *The semantic conception of truth and the foundation of semantics*, Philosophy and phemonenological research, Volume 4, 1944, Seite 341-375

[Tarski:1971] Alfred Tarski: *Der Wahrheitsbegriff in den formalisierten Sprachen*, Logik-Texte, Karel Berka & Lothar Kreiser, Akademie-Verlag, Berlin, 1971