

## Seminar: *Was ist Wahrheit?*

### 1. Blackboard & Kontakt

Alle Materialien für dieses Seminar finden Sie auf Blackboard, Passwort: The Truth. Sie erreichen mich per E-Mail: miguel.hoeltje@posteo.de.

### 2. Aktive Teilnahme

Die Anforderungen für die aktive Teilnahme umfassen zwei Bausteine:

#### 1. Posts im Blackboard-Forum

- Zu jedem Seminartext ist auf Blackboard ein Diskussionsforum angelegt. Zudem gibt es ein Forum, in welchem wir Themen für den Zwischenbilanztermin am 21.12. sammeln.
- Für die aktive Teilnahme ist es erforderlich, **zu mindestens 6 Sitzungen einen Post** im entsprechenden Diskussionsforum zu machen (gerne mehr). Ein Post als Vorbereitung des Zwischenbilanztermins ist für alle verpflichtend. Sie brauchen also **Zwischenbilanz + 5 weitere Posts**.
- Bitte posten Sie in den Foren zu den Seminartexten Fragen/Ideen/Vorschläge/Einwände zum jeweiligen Lektürestoff. Zudem können Sie sehr gerne auch auf bereits gepostete Fragen selbst antworten und/oder diese ergänzen.
- Am Zwischenbilanztermin ist Raum, auf offene Fragen zurückzukommen, Verbindungen zwischen besprochenen Themen herzustellen und den „größeren Rahmen“ zu thematisieren (Verbindungen zu anderen Philosophie-Bereichen und Disziplinen). Posten Sie hier entsprechend Ihre Themenwünsche und Diskussionsvorschläge.
- **Frist ist stets der Mittwoch vor der jeweiligen Sitzung.**

#### 2. Bearbeitung von Lektürefragen

- Zu jeder Sitzung, in der wir einen Seminartext besprechen, wird es ein paar kleine Lektürefragen geben.
- Für die aktive Teilnahme ist es erforderlich (zusätzlich zu Posts), **zu 2 Sitzungen die Lektürefragen zu bearbeiten** und die Bearbeitung über Blackboard abzugeben. (Posts und Lektürefragen-Bearbeitung sollten nicht zum selben Termin sein.)
- Die Lektürefragen lassen sich in der Regel kurz beantworten (**1-2 Seiten**); wer mag, kann aber gerne mehr schreiben! Ich werde alle Bearbeitungen lesen und ggf. in die Sitzungsvorbereitung einfließen lassen. Die Bearbeitungen bleiben unbenotet.
- Um die Teilnehmenden einigermaßen gleichmäßig zu verteilen, werden wir im Voraus festlegen, wer zu welchem Termin Lektürefragen bearbeiten wird. Näheres dazu demnächst.
- **Auch hier ist die Frist Mittwoch vor der jeweiligen Sitzung.**

Prinzipiell gibt es noch die Möglichkeit, als Alternative zur Bearbeitung der Lektürefragen zum Ende des Semester einen eigenen Aufsatz zu verfassen und diesen im Seminar zur Diskussion zu stellen:

**Alternative zu Lektürefragen: Aufsatz mit *Peer-Review***

- Am 15.02. gäbe es die Gelegenheit, einen Aufsatz aus Ihren Reihen zu besprechen.
- Dieser Aufsatz würde eine (im Rahmen des Seminarthemas) frei wählbare Fragestellung behandeln und von allen Teilnehmenden gelesen und diskutiert werden.
- Umfang: 2.000 bis 3.000 Wörter, Abgabe auf Blackboard bis 08.02.
- Sollten Sie diese Option wählen, besteht Ihre Aufgabe aus zwei Teilen:
  - Auswahl einer Fragestellung (in Rücksprache mit dem Seminarleiter)
  - rechtzeitiges Verfassen Ihres eigenen Aufsatzes
- Option 3 ist arbeitsintensiv, aber attraktiv:
  - tolle Gelegenheit, ausführliches Feedback zu einer eigenen Arbeit zu erhalten
  - ideale Vorbereitung auf eine mögliche Hausarbeit, die aus der Überarbeitung und Ausweitung des Aufsatzes hervorgehen könnte.
- Sollte es mehrere Interessent\*innen geben, losen wir.

### 3. Semesterüberblick

- 19.10. **Willkommen & Organisatorisches**
- 26.10. **Organisatorisches & Einführung**
- 02.11. **Gottlob Frege (1918): „Der Gedanke“ (Anfang bis S.204: „... von denen einige wahr, andere falsch sein können.“)**
- 09.11. **G. E. Moore (1911): „Wahre und falsche Meinungen“**
- 16.11. **Bertrand Russell (1912): „Wahrheit und Falschheit“**
- 23.11. **Die semantische Wahrheitskonzeption von Alfred Tarski (Handout folgt)**
- 30.11. **Chase B. Wrenn (2015): „Truth and Value“**
- 07.12. **Chase B. Wrenn (2015): „Epistemic Theories of Truth“**
- 14.12. **Frank P. Ramsey (1927/28): „Wahrheit: Ein Scheinproblem und das wirkliche Problem“**
- 21.12. **Zwischenbilanz**
- 11.01. **Paul Horwich (1998): „Postscript“, Abschnitte 1 bis 4**
- 18.01. **Crispin Wright (1995): „Wahrheit in der Ethik“, Abschnitte I und II**
- 25.01. **Crispin Wright (1995): „Wahrheit in der Ethik“, Abschnitte III und IV**
- 01.02. **Max Kölbel (2004): „Faultless Disagreement“, Abschnitte I-III**
- 08.02. **Max Kölbel (2004): „Faultless Disagreement“, Abschnitte IV-VII**
- 15.02. **Abschluss (& ggf. *Peer-Review*)**

### 4. Für die nächste Sitzung (26.10.)

- Zur Sitzung am 26.10. ist noch kein Text zu lesen.
- Stattdessen besteht Ihre Aufgabe darin, sich zunächst selbst Gedanken zur Leitfrage dieses Seminars zu machen: Was ist Wahrheit?
- Wie würden Sie selbst (noch vor der Lektüre von „Fachliteratur“) an diese Frage herangehen? Wie würde Sie selbst antworten? Gibt es Ihrer Auffassung nach Unklarheiten, die erst einmal

ausgeräumt werden müssen, bevor man eine Antwort entwickeln kann? Ist die Frage vielleicht mehrdeutig? Und was ist eigentlich die Relevanz dieser Frage? Wofür könnte eine Klärung wichtig sein?

- Die Sitzung am 26.10. werden wir unter anderem dazu nutzen, Ihre Vorschläge und Ideen zu sammeln, etwas zu systematisieren und gemeinsam zu diskutieren.
- Im Laufe des Semesters haben wir dann Gelegenheit, die eigenen ersten Überlegungen und Erwartungen mit den Argumenten und Thesen, die wir in den zu lesenden Texten finden werden, abzugleichen.

## 5. Zum Umgang mit den Seminartexten

- Die Texte sind oft schwierig (und teilweise unklar). Lassen Sie sich dadurch nicht entmutigen! Unser Anspruch ist nicht, alles zu verstehen; und schon gar nicht, alles sofort zu verstehen.
- In der Auseinandersetzung mit philosophischen Texten ist es bereits ein echter Fortschritt zu erkennen, *was* unklar ist und *wo* wir Verständnisprobleme haben.
- Das Kämpfen mit schwierigen Texten erfüllt einen Doppelzweck: Wir wollen etwas über die Materie lernen; und wir wollen uns generell darin üben, mit Texten zu arbeiten, Unklarheiten zu erkennen und mit anderen zu besprechen, etc.
- Die Auseinandersetzung mit philosophischen Texten kann Freude machen, aber auch zu Unsicherheit und Orientierungslosigkeit beitragen. Es dauert lange, in einem so alten Fach wie der Philosophie das Gefühl zu bekommen, man habe einen Überblick.
- Planen Sie hinreichend Zeit für die Lektüre ein und fangen Sie nicht zu spät damit an.

## 6. Ein paar Selbstverständlichkeiten zur Atmosphäre im Seminar

Lassen Sie uns für eine angenehme und produktive Atmosphäre sorgen. Ein paar Punkte:

- *Gleichzeitigkeit von Respekt, Wohlwollen und Kritikfähigkeit*: Diskussionen können hitzig werden. Das muss nicht schlecht und kann sogar gut sein. Wichtig ist, Inhalt und Person zu trennen. Das Seminar soll ein geschützter Raum für Personen sein (wir dürfen uns ausprobieren ohne Angst vor Fehlern), aber kein geschützter Raum für Inhalte (es ist erst einmal erlaubt, alles zu hinterfragen und auch Kritik zu üben). Stets respektvoll und wohlwollend, aber auch mal hart in der Sache.
- *Gewissenhafte Vorbereitung*: Das Seminar profitiert davon, dass alle gut vorbereitet sind. Wir lesen rechtzeitig die Texte, machen uns Gedanken, was diskutiert werden sollte, etc.
- *Arbeitsklima*: Pünktlicher Anfang, einander zuhören und ausreden lassen, möglichst wenig Ablenkungen. Ich empfehle, Texte oldschool in Papierform mitzubringen und auf Laptops und Tablets zu verzichten. Sollte das mal nicht möglich sein, ist ein behutsamer Einsatz wünschenswert (das Ablenkungspotential für Sie selbst und andere ist enorm): die Texte elektronisch lesen ist in Ordnung; mal eben eine E-Mail schreiben oder Social-Irgendwas checken sind es nicht. *Der Gebrauch von Handys/Smartphones sollte während der Sitzung bitte vollständig unterbleiben.*
- *Kleingesprächsrunden*: Wir werden weitgehend in großer Runde gemeinsam diskutieren. Aber oft ist es hilfreich, sich mal schnell mit den Nachbar:innen auszutauschen! Hierfür kann die gemeinsame Diskussion unterbrochen werden, damit für ein paar Minuten alle nach Lust und Laune mit allen anderen reden können (Fragen stellen, Antworten geben, Ideen sammeln). Sie können jederzeit anregen, dass jetzt ein guter Moment für eine solche Kleingesprächsrunde wäre!

# Tarskis semantische Wahrheitskonzeption

Miguel Hoeltje  
Freie Universität Berlin

Alfred Tarski (1901–1983) war ein aus Warschau gebürtiger Logiker, Mathematiker und Philosoph, der ab 1939 in den USA lebte und lehrte (von Mitte der 1940er Jahre bis zu seinem Tode an der *University of California*, Berkeley). Dieses Handout gibt eine Einführung in die Grundlagen von Tarskis Wahrheitskonzeption.

## 1. Hintergrund & Motivation

- Es liegt nahe, einige für die formalen Wissenschaften (Logik & Mathematik) zentrale Begriffe in Rekurs auf *Wahrheit* zu definieren.
- Beispielsweise wird der Begriff der *Folgerung* typischerweise mittels Wahrheit erläutert: Dass ein Satz/eine Formel  $S$  aus einer Menge  $\Phi$  von Sätzen/Formeln *folgt* heißt, dass  $S$  *wahr* sein muss sofern alle Elemente von  $\Phi$  *wahr* sind.
- Doch ohne weitere Erläuterungen mag man befürchten, dass die Rede von Wahrheit nicht mathematisch/logischen Standards von Klarheit und Präzision genügt.
- Insbesondere war seit langem bekannt, dass der Begriff der Wahrheit in eine Reihe von *Paradoxien* verwickelt ist. Betrachten Sie den Satz (L):
  - (L) Der neunte Satz auf dem Handout zu Tarskis Wahrheitsauffassung ist falsch.
- Noch bevor wir nachzählen können wir folgendes festhalten:
  - (a) Satz (L) ist wahr  $\leftrightarrow$  der neunte Satz ist falsch.
- Sobald wir nachzählen stellen wir fest:
  - (b) Satz (L) = der neunte Satz.
- Aus (a) und (b) ergibt sich:
  - (c) Satz (L) ist wahr  $\leftrightarrow$  Satz (L) ist falsch.
- Jetzt sind wir auf dem Weg in einen Widerspruch! So nehmen wir etwa in der klassischen Logik das Prinzip der *Bivalenz* an, demzufolge jeder Satz wahr oder falsch ist. Aus (c) ergibt sich also:
  - (d) Satz (L) ist wahr *und* Satz (L) ist falsch.
- Doch Bivalenz besagt ebenfalls, dass kein Satz *sowohl* wahr *als auch* falsch ist. Zusammen führen Bivalenz und (d) also in eine strikte Kontradiktion.
- Wir sind von (scheinbar?) harmlosen Annahmen mittels (scheinbar?) plausibler Schlussprinzipien in einen Widerspruch geraten. Das ist eine Paradoxie!
- Dies ist nur ein Beispiel aus einer großen Klasse verwandter Beispiele die zeigen, dass sich mittels eines „naiven“ Wahrheitsbegriffes Widersprüche herleiten lassen, ohne dass offensichtlich ist, an welcher Stelle genau etwas schiefgegangen ist.
- Tarskis Ziel bestand in erster Linie darin vorzuführen, wie sich für formalisierte Sprachen Wahrheitsdefinitionen geben lassen, die zum einen dem intuitiven Begriff von Wahrheit in Anwendung auf diese Sprachen gerecht werden, die aber zum anderen gleichzeitig formal präzise und frei von Paradoxien sind.
- Im Anschluss an Tarskis Arbeiten ist sein Ansatz auch auf natürliche Sprachen übertragen worden und bildet die Grundlage für eines der bis heute einflussreichsten Forschungsparadigma in der linguistischen Semantik und der Sprachphilosophie.

## 2. Ein paar Annahmen und Begriffe

- Tarskis Ansatz zielt auf *Sprachen* ab und als Wahrheitswertträger werden *Sätze* genommen (im Unterschied etwa zu *Gedanken/Propositionen* bei Frege oder *Meinungen* bei Moore und Russell).
- Eine Tarskische Wahrheitsdefinition besteht in der Definition eines Wahrheitsprädikates für eine Sprache. Für eine gegebene Sprache  $\mathcal{L}$  werden wir als Wahrheitsprädikat „ist wahr in  $\mathcal{L}$ “ verwenden.
- Die Sprache, für die ein Wahrheitsprädikat definiert wird, nennen wir *Objektsprache*, jene, in der die Definition formuliert ist, *Metasprache*.
- Damit eine Definition von „ist wahr in  $\mathcal{L}$ “ angemessen ist, muss das definierte Prädikat auf genau die wahren Sätze der Objektsprache  $\mathcal{L}$  zutreffen.
- Tarskis Grundidee besteht in der Entwicklung eines Kriteriums, dessen Erfüllung dies sicherstellen soll: *Konvention W*.

## 3. Tarskis Grundidee: *Konvention W*

- Zur Illustration nehmen wir als Objektsprache Englisch ( $\mathcal{L}_E$ ) und als Metasprache Deutsch und betrachten ein Beispiel:
    - (1) There are intelligent alien life forms.
  - Bei der Definition von „ist wahr in  $\mathcal{L}_E$ “ müssen wir sicherstellen, dass dieses Prädikat auf (1) zutrifft, sofern (1) wahr ist, und dass es nicht auf (1) zutrifft, sofern (1) nicht wahr ist.
  - Doch wie können wir dies sicherstellen, wenn wir den Wahrheitswert von (1) gar nicht kennen?
  - Da wir  $\mathcal{L}_E$  verstehen, können wir unser Wissen darüber, was dieser Satz bedeutet heranziehen: Wenn es intelligente außerirdische Lebensformen gibt, so ist (1) wahr; wenn nicht, dann nicht:
- W1 „There are intelligent alien life forms“ ist wahr in  $\mathcal{L}_E \leftrightarrow$  es gibt intelligente außerirdische Lebensformen.
- W1 gibt die *Wahrheitsbedingungen* von (1) an. Wir nennen solche Sätze *W-Sätze*:

### W-Satz

Ein W-Satz ist ein Satz der folgenden Form:  $S$  ist wahr in  $\mathcal{L} \leftrightarrow p$ .

- Ob (1) wahr ist oder nicht: sofern eine Definition von „ist wahr in  $\mathcal{L}_E$ “ den W-Satz W1 impliziert, ist sichergestellt, dass diese Definition den Beispielsatz (1) korrekt behandelt.
- Denn eine solche Definition knüpft das Zutreffen des Wahrheitsprädikates auf Satz (1) an eine Bedingung, die tatsächlich notwendig und hinreichend für die Wahrheit von (1) ist: die Existenz intelligenter außerirdischer Lebensformen.
- Dass mit W1 das Zutreffen von „ist wahr in  $\mathcal{L}_E$ “ tatsächlich an die korrekte Bedingung geknüpft wird, wurde dadurch garantiert, dass wir auf der rechten Seite des Bikonditionals „ $\leftrightarrow$ “ einen Satz der Metasprache verwendet haben, der mit dem fraglichen Satz der Objektsprache *synonym* ist.
- Solche W-Sätze nennen wir *interpretierend*:

### Interpretierender W-Satz

Ein W-Satz „ $S$  ist wahr in  $\mathcal{L} \leftrightarrow p$ “ ist *interpretierend* wenn  $S$  und „ $p$ “ dieselbe Bedeutung haben.

- Tarskis Adäquatheitskriterium für Wahrheitsdefinitionen ergibt sich aus der Verallgemeinerung des Gedankengangs, den wir gerade am Beispiel (1) nachvollzogen haben.
- Kurz gefasst fordert dieses Adäquatheitskriterium, dass eine Wahrheitsdefinition interpretierende W-Sätze implizieren soll.
- Die vollständige Fassung des Kriteriums, welches Tarski *Konvention W* nennt, lautet:

**Konvention W**

Eine in einer Metasprache  $\mathcal{ML}$  formulierte und formal korrekte (d.h. konsistente und nicht-zirkuläre) Definition eines Wahrheitsprädikates „ist wahr in  $\mathcal{L}$ “ für eine Objektsprache  $\mathcal{L}$  ist genau dann *inhaltlich adäquat*, wenn sie alle Instanzen des Schemas

$$S \text{ ist wahr in } \mathcal{L} \leftrightarrow p$$

impliziert, die sich erzeugen lassen, indem „ $S$ “ durch eine Bezeichnung eines  $\mathcal{L}$ -Satzes ersetzt wird und „ $p$ “ durch einen synonymen Satz der Metasprache  $\mathcal{ML}$ .

- Wir werden uns nun anhand einiger einfacher Beispielsprachen anschauen, wie sich Definitionen geben lassen, die Konvention W erfüllen.

#### 4. Beispielsprachen & Beispieldefinitionen

##### Eine endliche „Sprache“

- Werfen wir zunächst einen Blick auf *endliche* Sprachen, also Sprachen, die nur eine endliche Anzahl von Sätzen enthalten.
- Im Fall von endlichen Sprachen ist es vergleichsweise einfach, eine Wahrheitsdefinition anzugeben, die Konvention W erfüllt.
- Als Beispiel betrachten wir  $\mathcal{L}_1$  und nehmen an, dass diese Sprache nur zwei Sätze enthält: „There are intelligent alien life forms“ und „Socrates is human“.
- Hier ist eine Definition von „ist wahr in  $\mathcal{L}_1$ “ die Konvention W erfüllt:

**Definition 1**

Ein beliebiger Satz  $S$  ist wahr in  $\mathcal{L}_1$  genau dann, wenn gilt:

$S = \text{"There are intelligent alien life forms"}$  und es gibt intelligente außerirdische Lebensformen

oder

$S = \text{"Socrates is human"}$  und Sokrates ist ein Mensch.

- Oftmals ist es übersichtlicher anstelle einer expliziten Definition eine *Theorie* (verstanden als eine Menge von Axiomen) anzugeben, aus der sich die gewünschten interpretierenden W-Sätze herleiten lassen.
- Im Fall einer endlichen Sprache kann eine solche Theorie eine recht triviale Form annehmen, da wir schlicht die gewünschten interpretierenden W-Sätze direkt als Axiome nehmen können:

**Theorie 1**

W1 „There are intelligent alien life forms“ ist wahr in  $\mathcal{L}_1 \leftrightarrow$  es gibt intelligente außerirdische Lebensformen.

W2 „Socrates is human“ ist wahr in  $\mathcal{L}_1 \leftrightarrow$  Sokrates ist ein Mensch.

- Dieses Vorgehen ist offensichtlich unabhängig davon, wie viele Sätze die fragliche Sprache enthält (solang es nur endlich viele sind): für jeden weiteren Satz könnten wir Definition 1 ein weiteres Disjunkt oder Theorie 1 ein weiteres Axiom hinzufügen.
- Solange die Anzahl der Sätze endlich ist, lässt sich nach diesem Vorgehen also stets eine Wahrheitsdefinition (bzw. eine entsprechende Theorie) bilden.
- Interessanter wird es, wenn wir uns *unendlichen* Sprachen zuwenden.

### Eine satzlogische Sprache

- Betrachten wir nun eine Erweiterung von  $\mathcal{L}_1$  um den Negationsoperator „it is not the case that“ und den Konjunktiv „and“, und nennen wir die so erweiterte Sprache  $\mathcal{L}_2$ .
- In  $\mathcal{L}_2$  lassen sich nun neue Sätze formulieren, unter anderem:
  - (2) It is not the case that there are intelligent alien life forms.
  - (3) There are intelligent alien life forms and Socrates is human.
  - (4) There are intelligent alien life forms and it is not the case that Socrates is human.
- Satzoperatoren wie „it is not the case that“ und „and“ lassen sich wieder auf die mit ihrer Hilfe gebildeten Sätze anwenden. So können wir, ausgehend von (2), durch eine weitere Anwendung von „it is not the case that“ einen weiteren Satz bilden:
  - (5) It is not the case that it is not the case that there are intelligent alien life forms.
- Da sich solche Anwendungen beliebig häufig wiederholen lassen, umfassen Sprachen wie  $\mathcal{L}_2$  *unendlich viele* Sätze (analog für den Konjunktiv „and“).
- Das Vorgehen, welches wir im Fall der endlichen Sprache  $\mathcal{L}_1$  für die Formulierung einer Wahrheitsdefinition verwendet haben lässt sich auf unendliche Sprachen nicht einfach übertragen, da die entsprechende „Definition“ unendlich lang sein müsste (bzw. die entsprechende Theorie unendlich viele Axiome enthalten würde).
- Stattdessen werden wir für  $\mathcal{L}_2$  eine *rekursive Definition* konstruieren.
- In einer solchen rekursiven Definition geben wir zunächst (in den sogenannten *Basisklauseln*) für die einfachen Sätze direkt an, unter welchen Bedingungen sie wahr sind.
- Im Anschluss geben wir in den *rekursiven Klauseln* an, wie sich die Wahrheitsbedingungen für Sätze, die mittels der Satzoperatoren gebildet wurden, aus den Wahrheitsbedingungen der enthaltenen Sätze ergeben.

#### Definition 2

Ein beliebiger Satz  $S_1$  ist wahr in  $\mathcal{L}_2$  genau dann, wenn gilt:

$S_1 = \text{„There are intelligent alien life forms“}$  und es gibt intelligente außerirdische Lebensformen

oder

$S_1 = \text{„Socrates is human“}$  und Sokrates ist ein Mensch.

oder

$S_1$  ist das Ergebnis der Verknüpfung von „it is not the case that“ mit einem Satz  $S_2$  und es ist nicht der Fall, dass  $S_2$  wahr in  $\mathcal{L}_2$  ist.

oder

$S_1$  ist das Ergebnis der Verknüpfung zweier Sätze  $S_2$  und  $S_3$  mittels „and“ und  $S_2$  ist wahr in  $\mathcal{L}_2$  und  $S_3$  ist wahr in  $\mathcal{L}_2$ .

- Hier nochmal im Theorie-Format und mit ein wenig formallogischer Notation:

**Theorie 2**

- W1 „There are intelligent alien life forms“ ist wahr in  $\mathcal{L}_2 \leftrightarrow$  es gibt intelligente außerirdische Lebensformen.
- W2 „Socrates is human“ ist wahr in  $\mathcal{L}_2 \leftrightarrow$  Sokrates ist ein Mensch.
- N  $\forall S$  („it is not the case that“ $\neg S$  ist wahr in  $\mathcal{L}_2 \leftrightarrow \neg(S$  ist wahr in  $\mathcal{L}_2)$ )
- K  $\forall S_1 \forall S_2 (S_1 \wedge S_2$  ist wahr in  $\mathcal{L}_2 \leftrightarrow (S_1$  ist wahr in  $\mathcal{L}_2 \wedge S_2$  ist wahr in  $\mathcal{L}_2))$

- Die Grundidee bei dieser Behandlung von Satzoperatoren besteht also darin, den Beitrag, den ein Satzoperator zu den Wahrheitsbedingungen eines ihn enthaltenden Satzes leistet zu spezifizieren, indem wir in der Metasprache einen gleichbedeutenden Satzoperator auf Wahrheitszuschreibungen für die im Ausgangssatz enthaltenen Sätze anwenden.
- Auf diese Weise können (zumindest wahrheitsfunktionale) Satzoperatoren auf kompositionale Weise behandelt werden.

**Namen und (einstellige) Prädikate**

- Bislang haben wir für den Satz „Socrates is human“ *direkt* die Wahrheitsbedingungen angegeben.
- Doch dieser Satz besteht ja eigentlich aus mehreren Ausdrücken die jeweils eigene semantische Eigenschaften haben, wobei sich die Wahrheitsbedingungen des Satzes aus dem Zusammenspiel dieser semantischen Eigenschaften ergeben.
- Grob gesprochen: Dass „Socrates is human“ genau dann wahr ist, wenn Sokrates ein Mensch ist, ergibt sich daraus, dass der Name „Socrates“ einen bestimmten Gegenstand bezeichnet (nämlich Sokrates), und dass mit dem Prädikat „is human“ eine bestimmte Eigenschaft zugeschrieben wird (nämlich die, ein Mensch zu sein).
- Wir können nun eine Wahrheitsdefinition bzw. eine Wahrheitstheorie formulieren, welche diese subsententialen Ausdrücke als Ausgangspunkt nimmt.
- Hierzu werden wir von zwei neuen Begriffen Gebrauch machen:
  - Wir werden den von einem Name  $N$  in einer Sprache bezeichneten Gegenstand den *Bezugsgegenstand* von  $N$  nennen und dies als  $B(N)$  abkürzen. Es gilt also beispielsweise:  $B(\text{„Socrates“}) = \text{Sokrates}$ .
  - Für das einstellige Prädikat „is human“ geben wir an, unter welchen Bedingungen ein Gegenstand dieses Prädikat *erfüllt*; nämlich genau dann, wenn es sich bei dem Gegenstand um einen Menschen handelt.
- Aus Gründen der Übersichtlichkeit:
  - Wir schauen uns nicht die explizite Wahrheitsdefinition an, sondern beschränken uns auf das Theorie-Format.
  - Wir lassen den Bezug auf die Objektsprache implizit; so schreiben wir beispielsweise anstelle von „ist wahr in  $\mathcal{L}_2$ “ und „erfüllt in  $\mathcal{L}_2$ “ schlüssig: „ist wahr“ und „erfüllt“.
  - Die Axiome für Negation und Konjunktion listen wir nicht erneut auf.
  - Den komplizierteren Satz „There are intelligent alien life forms“ ignorieren wir erst einmal.

**Theorie 3**

- S       $B(\text{``Socrates''}) = \text{Sokrates}$ .
- H       $\forall x (x \text{ erfüllt ``is human''} \leftrightarrow x \text{ ist ein Mensch})$
- NP     Ein Satz, der aus einem Namen  $N$  gefolgt von einem einstelligen Prädikat  $P$  besteht ist genau dann wahr, wenn der Bezugsgegenstand von  $N$  das Prädikat  $P$  erfüllt.  
Mit etwas formallogischer Notation:  
 $\forall N \forall P (N \cap P \text{ ist wahr} \leftrightarrow B(N) \text{ erfüllt } P)$ .

- Aus den Axiomen von Theorie 3 können wir nun den folgenden interpretierenden W-Satz herleiten:

W2     „Socrates is human“ ist wahr  $\leftrightarrow$  Sokrates ist ein Mensch.

**Mehrstellige Prädikate & unendliche Folgen**

- Wir hatten oben das einstellige Prädikat „is human“ mittels des Begriffes der Erfüllung behandelt; ein Gegenstand erfüllt dieses Prädikat, wenn er ein Mensch ist.
- Doch natürlich gibt es auch *mehrstellige* Prädikate, beispielsweise „is taller than“ (zweistellige) oder „... is between ... and ...“ (dreistellig).
- Für die Behandlung mehrstelliger Prädikate müssen wir den Begriff der Erfüllung erweitern, da es hier keinen Sinn macht davon zu sprechen, dass sie von einem *einzelnen* Gegenstand erfüllt werden.
- Das zweistellige „is taller than“ beispielsweise drückt eine Relation aus, die zwischen zwei Gegenständen bestehen kann.
- Um die Wahrheitsbedingungen von Sätzen wie „Plato is taller than Socrates“ auf die Erfüllungsbedingungen des enthaltenen Prädikates zurückzuführen müssen wir zudem die *Reihenfolge* der Gegenstände beachten.
- Wir könnten hierzu von *geordneten Paaren* Gebrauch machen: (Plato, Sokrates) ist das geordnete Paar, welches Plato als erstes Element und Sokrates als zweites Element enthält. Wir könnten jetzt sagen, dass ein geordnetes Paar das Prädikat „is taller than“ genau dann erfüllt, wenn das erste Element des Paares größer ist als das zweite Element.
- Doch für dreistellige Prädikate bräuchten wir dann schon geordnete Tripel; für vierstellige Prädikate 4er-Tupel; und so weiter.
- Um auf einen Schlag Prädikate mit beliebiger Stelligkeit zu behandeln, können wir *unendliche Folgen* von Gegenständen verwenden (kurz: *Folgen*).
- Die Verwendung von Folgen von Gegenständen ist zudem vor allem für die Behandlung quantifizierter Sprache nützlich, die *Variablen* und (beliebig komplexe) *offene Sätze* enthalten.

**Variablen, offene Sätze und Quantoren**

- Aus Platzgründen gehen wir nicht ausführlich auf die Behandlung quantifizierter Sprachen ein.
- Aber hier zumindest ein schneller Blick auf eine Art, wie man nach Tarski Methode eine Wahrheitstheorie für eine solche Sprache formulieren kann:

**Theorie 4**

- W  $\forall S (S \text{ ist wahr} \leftrightarrow \forall \Gamma (\Gamma \text{ erfüllt } S))$ .
- H  $\forall v_i \forall \Gamma (\Gamma \text{ erfüllt } v_i \cap \text{"is human"} \leftrightarrow \text{das } i\text{-te Element von } \Gamma \text{ ist ein Mensch})$
- T  $\forall v_i \forall v_j \forall \Gamma (\Gamma \text{ erfüllt } v_i \cap \text{"is taller than"} \cap v_j \leftrightarrow \text{das } i\text{-te Element von } \Gamma \text{ ist größer als das } j\text{-te Element von } \Gamma)$
- N  $\forall \Phi \forall \Gamma (\Gamma \text{ erfüllt } \text{"it is not the case that"} \cap \Phi \leftrightarrow \neg(\Gamma \text{ erfüllt } \Phi))$ .
- K  $\forall \Phi_1 \forall \Phi_2 \forall \Gamma (\Gamma \text{ erfüllt } \text{"} \cap \Phi_1 \cap \text{"and"} \cap \Phi_2 \cap \text{"} \leftrightarrow (\Gamma \text{ erfüllt } \Phi_1 \wedge \Gamma \text{ erfüllt } \Phi_2))$ .
- E  $\forall \Phi \forall v_i \forall \Gamma_1 (\Gamma_1 \text{ erfüllt } \text{"every"} \cap v_i \cap \Phi \leftrightarrow \forall \Gamma_2 (\Gamma_2 \text{ ist eine } v_i\text{-Variante von } \Gamma_1 \rightarrow \Gamma_2 \text{ erfüllt } \Phi))$ .

**5. Zurück zu den Paradoxien**

- Teil der Motivation für Tarskis Ansatz bestand darin vorzuführen, wie sich Wahrheitsprädikate für formalisierte Sprachen präzise definieren lassen, ohne dass dabei Paradoxien entstehen.
- Für diesen Punkt ist die Trennung von Objektsprache und Metasprache zentral, sowie die Anforderungen, die Tarski an diese Sprache stellt.
- Kehren wir kurz zu dem problematischen Satz (L) zurück:

(L) Der neunte Satz auf dem Handout zu Tarskis Wahrheitsauffassung ist falsch.

- Dass wir einen Satz wie (L) im Deutschen formulieren können liegt daran, dass natürliche Sprachen wie das Deutsche (oder auch das Englische und andere natürliche Sprachen), wie Tarski es nennt, „semantically universal“ sind:

*Together with the linguistic objects, such as sentences and terms, which are components of this language, names of these objects are also included in the language (as we know, names of expressions can be obtained by putting the expressions in quotes); in addition, the language contains semantic terms such as “truth”, “name”, “designation”, which directly or indirectly refer to the relationship between linguistic objects and what is expressed by them.*

- Solche semantisch universalen Sprachen enthalten also sowohl das Material, um auf die Ausdrücke der eigenen Sprache Bezug zu nehmen, als auch das Material, um diese Ausdrücke semantisch zu charakterisieren.
- Insbesondere enthalten solche Sprachen das eigene Wahrheitsprädikat.
- Dieser „universale“ Charakter führt zu der Formulierbarkeit der problematischen paradoxen Sätze.
- Wenn wir uns nun auf Sprachen beschränken, die nicht ihr eigenes Wahrheitsprädikat enthalten, so ist in diesen Sprachen beispielsweise (L) nicht mehr formulierbar.
- Wir können dennoch auch für eine solche Sprache eine Wahrheitsdefinition angeben, indem wir eine Metasprache verwenden, die *reichhaltiger* als die Objektsprache ist; insbesondere wird diese Metasprache ein Wahrheitsprädikat für die Objektsprache enthalten (nicht aber das eigene Wahrheitsprädikat).
- Etwas verkürzt: Auf diese Weise können wir nun eine *Hierarchie* immer reichhaltigerer Sprachen konstruieren, in welcher die Sprache auf der jeweils „höheren“ Stufe dazu dient, Wahrheit für die Sprache auf der Stufe darunter zu definieren.
- Die Formulierbarkeit der betrachteten paradoxen Sätze wird also durch eine Einschränkung des „erlaubten“ Vokabulars verhindert.

- Gleichzeitig wird der Verlust an Ausdrucksstärke zumindest zu einem gewissen Grade durch die Einbettung in eine Sprachhierarchie aufgefangen, in welcher die Paradoxie-freie semantische Charakterisierung einer gegebenen Sprache mittels der nächst höheren Metasprache möglich ist.

## 6. Anwendung auf natürliche Sprachen

- Tarski hat Konvention W mit Blick auf formale Sprachen formuliert, und er selbst war skeptisch, dass sich seine Methode fruchtbar auf natürliche Sprachen anwenden lasse.
- Im Anschluss an Tarskis Arbeiten ist sein Ansatz auch auf natürliche Sprachen übertragen worden und bildet die Grundlage für eines der bis heute einflussreichsten Forschungsparadigma in der linguistischen Semantik und der Sprachphilosophie: Die sogenannte wahrheitskonditionale Semantik.
- Die Ausweitung dieses Ansatzes auf natürliche Sprachen steht u.a. vor den folgenden Herausforderungen:
  - Nicht-deklarative Sätze
  - Mehrdeutigkeit
  - Kontextabhängigkeit
  - Vagheit
  - Mögliche Wahrheitswertlosigkeit
  - Nicht-extensionale Konstruktionen
- Der Versuch, formal präzise semantische Theorien für natürliche Sprachen zu formulieren, die mit diesen (und einigen weiteren) Phänomenen adäquat umgehen können, bildet heute ein großes und lebhaftes Forschungsfeld im Schnittbereich zwischen Philosophie, Linguistik, Logik und Psychologie.

## Theorie/Analyse vs. Test/Kriterium

Hinsichtlich eines Begriffes *B* können wir 2 Fragen unterscheiden:

**F1** Wie lässt sich *B* analysieren/erklären? Was heißt es, *B* zu sein?

**F2** Wie können wir herausfinden, ob etwas *B* ist?

F1 ist begrifflich/metaphysisch und fragt nach einer Theorie/Analyse von *B*; F2 ist erkenntnistheoretisch und fragt nach einem Test/Kriterium für *B*.

F1 und F2 sind verschieden, aber nicht zusammenhanglos.

*Beispiel: B = hat Covid-19*

F1 Covid-19 zu haben heißt, durch eine Infektion mit dem SARS-CoV-2 Virus verursachte Krankheitssymptome zu haben.

F2 z.B. Anamnese plus PCR-Test.

Wenn im Folgenden von *Theorien* die Rede ist, so handelt es sich um Antwortversuche auf F1 für *Wahrheit*.

## Wahrheitswertträger (WT)

Theorien können sich u.a. darin unterscheiden, auf welche Arten von Gegenständen sie als WT abzielen. Gängige Optionen:

- Mentale Zustände (Moore, Russell, Ramsey)
- Sätze (Tarski)
- Propositionen (Freges „Gedanken“)

Zuweilen ist von *primären* und *sekundären* WT die Rede. *Beispiel:* Wer meint, dass ein wahrer Satz nur wahr ist, weil die durch den Satz ausgedrückte Proposition wahr ist, nimmt Propositionen als *primäre*, Sätze als *sekundäre* WT.

## Propositionen

Propositionen spielen in vielen Bereichen der Philosophie eine Rolle; aber sowohl die Existenz als auch die zentralen Eigenschaften von Propositionen sind umstritten. Wer Propositionen annimmt geht typischerweise davon aus, dass diese die folgenden Rollen spielen; Propositionen sind:

- Inhalte von mentalen Zuständen wie Überzeugungen und Wünschen
- Inhalte von Sprechakten wie Behauptungen und Satzfragen
- Träger von Wahrheitswerten und deren Modalitäten (notwendig, contingenter, etc.) sowie die Relata von Beziehungen wie Folgerung
- das, was Deklarativsätze (relativ zu einem Äußerungskontext) ausdrücken.

## Über Wahrheit sprechen

### Wahrheitsprädikat vs. Wahrheitsoperator

Wir können verschiedene Konstruktionen verwenden, um Wahrheit zuzuschreiben. Der Unterschied zwischen einem Prädikat und einem Satzoperator besteht darin, welche Art von Ausdruck als Input dient: Bei Prädikaten sind es *singuläre Terme*, bei Satzoperatoren sind es *Sätze*.

Prädikat: Der Satz „Snow is white“ ist wahr. // Ihre letzte Behauptung ist wahr.

Operator: Es ist wahr, dass Schnee weiß ist.

### Explizit gegeben vs. beschrieben (⇒ Ramsey)

Explizit: Es ist wahr, dass Schnee weiß ist.

Beschrieben: Was sie gesagt hat, ist wahr. // Alles, was der Papst sagt, ist wahr.

### Verhältnis „einfache Sätze“/Wahrheitszuschreibungen

In welchem Verhältnis stehen (1) und (1<sub>O</sub>)/(1<sub>P</sub>)?

(1) Die Erde ist rund.

(1<sub>O</sub>) Es ist wahr, dass die Erde rund ist.

(1<sub>P</sub>) „Die Erde ist rund“ ist wahr.

### Die Redundanzthese

ist eine These über den Operator „Es ist wahr, dass“. Sie besagt, dass mit „Es ist wahr, dass *p*“ genau dasselbe gesagt wird wie mit „*p*“. Dieser These zufolge ist die Rede von Wahrheit in solchen Fällen redundant: man kann den Operator schlicht weglassen. (⇒ Ramsey)

### Die Omnipräsenzthese

ist eng verwandt mit der Redundanzthese, betont aber einen anderen Aspekt. Der Omnipräsenzthese zufolge ist der Begriff der Wahrheit „allgegenwärtig“: Da mit „Es ist wahr, dass *p*“ genau dasselbe gesagt wird wie mit „*p*“ ist der Begriff der Wahrheit in jeder Aussage bereits enthalten. (⇒ Frege)

**Substanzelle Theorien**

Wahrheit ist eine substanzelle Eigenschaft,  
deuten Wesen eine W-Theorie angeben sollte

**Korrespondenztheorie**

- ! Wahrheit ist Übereinstimmung mit „der Welt“. Varianten ergeben sich durch Wahl der Relata: Was sind WTs? Mit was können diese übereinstimmen? Beispiele Moore & Russell: wahre *Meinungen* stimmen mit *Tatsachen* überein.
- ? Was ist *Übereinstimmung*? Was sind *Tatsachen*? Übertragung auf *Falschheit*?

**Kohärenztheorie**

- ! Wahrheit kommt WTs nicht unabhängig von anderen WTs zu, sondern besteht in der Einbettung in ein kohärentes (≈ konsistent & explanatorisch verknüpft) & umfassendes (≈ beantwortet alle Fragen) System.
- ? Existenz mehrerer kohärenter & umfassender Systeme, die miteinander inkompatibel sind.  
Könnte es nicht auch gänzlich *unwissbare* Wahrheiten geben?

**Pragmatismus**

- ! James: Wahr ist, was nützlich ist/was funktioniert.  
Peirce: Wahr ist, was alle am idealen Ende aller Untersuchungen akzeptieren würden.
- ? Falsches kann sehr nützlich sein/funktionieren.  
Was heißt *ideales Ende*? Können wir das ohne Wahrheit erklären?  
Warum sollte sich hinsichtlich *Am 20.12.2023 war die Anzahl der Haare auf meinem Kopf gerade am „idealen Ende“ ein Konsens ergeben?*  
Könnte es nicht auch gänzlich *unwissbare* Wahrheiten geben?

**Epistemische Theorien**  
Wahrheit ist mittels epistemischer Begriffe definiert  
und/oder Wahrheiten sind stets wissbar**Tarskis semantische Konzeption**

- ! Zielt auf *Sätze formalisierter Sprachen* als WT.  
Unterscheidet *Objektsprache* & *Metasprache* (u.a. zur Vermeidung von Paradoxien)
- Entwickelt ein Adäquatheitskriterium für *Definitionen* von *Wahrheitsprädikaten* (*Konvention W*): Diese sollten interpretierende Theoreme der Form „*S* ist wahr  $\leftrightarrow p$ “ implizieren.
- ? Ausweitung auf natürliche Sprachen?  
Was ist mit WTs anderer Arten (mentale Zustände und/oder Propositionen)?

**Die Redundanztheorie**

geht noch über die Redundanzthese hinaus und sagt, dass die Rede von Wahrheit in *allen* Fällen redundant ist, auch wenn wir in einigen Fällen gezwungen sind, aus grammatischen Gründen beispielsweise „Es ist wahr, dass“ zu verwenden. ( $\Rightarrow$  Explizit gegeben vs. beschrieben). Dieser Theorie zufolge gibt es kein interessantes „Wesen“ der Wahrheit, das eine philosophische Theorie ausbuchstabieren könnte; Wahrheit erfüllt eine rein „grammatische Funktion“.

**Ramseys Theorie**

- ! M ist wahr  $\leftrightarrow \exists p [(M \text{ ist eine Meinung, dass } p) \wedge p]$
- ? Wie genau funktioniert die Quantifikation in Satzposition?  
Was heißt es, *eine Meinung, dass p* zu sein?

**Minimalismus**

Als nächstes im Seminar: Der wahrheitstheoretische Minimalismus von Paul Horwich ist ein weiteres Beispiel für einen „deflationären“ Ansatz.

**Frege zur Undefinierbarkeit von Wahrheit**

Ausgehend von der Omnipräsenzthese präsentiert Frege ein Argument dafür, dass Wahrheit nicht definierbar ist. (Scheint etwas mit einem unendlichen Regress oder einem Zirkelvorwurf zu tun zu haben; die genaue Form wird kontrovers diskutiert.)

Wohlgemerkt: Frege meint *nicht*, Wahrheit sei irgendwie „defekt“ oder „ungenau“ oder, dass man nie herausbekommen könnte, ob etwas wahr sei. Vielmehr meint Frege, dass Wahrheit ein basaler, unanalysierbarer Begriff ist.

**Deflationäre Theorien**  
Wahrheit ist nicht substanzell. W-Theorien sind „trivial“

## Wahrheit und Wert

Im gelesenen Abschnitt aus Wrenns Kapitel „Truth and Value“ werden 2 mögliche Thesen über den Wert der Wahrheit diskutiert:

- (G) Wahrheit ist wesentlich eine „kind of goodness“.
- (I) Wahrheit ist intrinsisch wertvoll.

### Dummett und (G)

Ein Autor, den Wrenn im Zusammenhang mit (G) diskutiert, ist Dummett. Wrenn interpretiert Dummetts Ausführungen so, dass im Hintergrund das folgende Argument für (G) steht:

- (1) Wenn es zum Begriff der Wahrheit gehört, dass wahre Behauptungen besser sind als falsche, dann ist Wahrheit eine normative Eigenschaft.
- (2) Es gehört zum Begriff der Wahrheit, dass wahre Behauptungen besser sind als falsche.
- (3) Also: (G) Wahrheit ist wesentlich eine „kind of goodness“.

Im Dummett-Zitat auf S.37f. interpretiert Wrenn dann als ein Versuch, (2) zu begründen. In diesem Zusammenhang spielt Dummetts Analogie des Schachspiels und der Rolle des Mattsetzens eine zentrale Rolle.

### Lynchs Gedankenexperiment/Argument für (I)

Ist Wahrheit intrinsisch wertvoll? Ist etwas schon allein deshalb wertvoll, weil es wahr ist? Wrenn referiert ein Gedankenexperiment von Lynch, aus welchem sich ein Argument für (I) ergeben soll. Wrenn selbst lehnt dieses Argument dann aber als nicht überzeugend ab.

*Gedankenexperiment:* Sie haben die Wahl zwischen 2 Optionen.

*Option A:* Normales Leben.

*Option B:* Matrix-Szenario (für Sie fühlt sich alles genauso an wie in Option A, aber Sie leben in einer Computersimulation).

Nehmen wir an, Sie würden sich lieber für A entscheiden.

*Argument:*

- (1) Der einzige relevante Unterschied zwischen A und B besteht darin, dass Sie in A mehr wahre Überzeugungen haben.
- (2) Also scheinen Sie zu meinen, dass Wahrheit intrinsisch wertvoll ist.
- (3) Sie sind gut darin zu erkennen, ob etwas intrinsisch wertvoll ist.
- (4) Also: (I) Wahrheit ist intrinsisch wertvoll.

*Wrenns Kritik an diesem Argument:* Annahme (1) ist inkorrekt.

A und B unterscheiden sich beispielsweise auch darin, welche Ihrer Wünsche und Hoffnungen erfüllt werden. Wenn Sie beispielsweise einer kranken Person Beistand leisten möchten, so wird dieser Wunsch in B enttäuscht (dort kommt es ihnen nur so vor, als würden Sie Beistand leisten).

## Leitfaden zu Hausarbeiten

Hausarbeiten sollten sich an dem entsprechenden Leitfaden des Instituts für Philosophie orientieren.

Ein paar weitere Hinweise finden Sie zudem in dem entsprechenden Handout (bald auf Blackboard).

### Absprache Hausarbeiten

Die Absprache einer Hausarbeit sollte sich in 5 Schritte aufteilen:

1. Vorschlag für Textgrundlage, Fragestellung, These & Abgabedatum
2. Rückmeldung von mir zu diesem Vorschlag
3. Abgabe eines Exposés für die Hausarbeit

Ein Exposé dient dazu, noch vor dem Verfassen der Hausarbeit ein gezieltes Feedback zu ermöglichen.

Aus dem Exposé sollte hierfür hervorgehen: Was ist die geplante primäre Textgrundlage? Welche weiteren Texte sollen voraussichtlich zudem behandelt werden? Was ist die leitende Fragestellung? Was wird die zu untersuchende These sein? Was ist die geplante Struktur der Arbeit? Was soll (grob) in den geplanten Abschnitte der Arbeit passieren?

Eine ausführlichere Erläuterung zu Exposés für Hausarbeiten entnehmen Sie dem entsprechenden Handout (bald auf Blackboard).

4. Rückmeldung von mir zum Exposé
5. Auf Basis der Rückmeldung und in Rücksprache mit Ihnen endgültige Festlegung von Textgrundlage etc.