EPISTEMISCHE LOGIK
Grundlagen

Unwissbarkeit

Manuel Bremer
University of Düsseldorf, Germany
www.mbph.de

Eigenes Unwissen

• Auch wenn es *de facto* so ist, dass ich eine Wahrheit nicht weiß, kann ich nicht von einer konkreten Wahrheit wissen, dass ich diese nicht weiß:

$$W_i(t \wedge \neg W_i t)$$

ist inkonsistent. [Warum?]

• Wenn ich von einer konkreten Wahrheit wissen wollte, dass ich sie nicht weiß, muss ich sie als Wahrheit kennen, also an sie glauben, also sie wissen ↯

Logisch ergibt sich die Inkonsistenz durch [∧Dis].

• Konsistent hingegen ist:

$$\vdash W_i(\exists t \neg W_i t)$$

denn Existentielle Spezialisierung führt nicht zu einer konkreten Wahrheit, die ich nicht kenne. (Es handelt sich sogar um ein Theorem, dass aus der Kenntnis des Axioms der Nichtallwissenheit folgt.)

Mit $W_i(\exists t \neg W_i t)$ weiß ich sofort $W_i(\exists t \neg G_i t)$, denn falsch kann eine Wahrheit nicht sein.

- Solange ich etwas nur für möglich halte, kann ich glauben, dass ich es nicht weiß – und dass sogar niemand es weiß.

- Alle folgenden Sätze sind konsistent:

$M_i(p \wedge \neg W_i p)$	und	$M_i(p \wedge \forall a \neg W_a p)$
$G_i(\Diamond p \wedge \neg W_i p)$	und	$G_i(\Diamond p \wedge \forall a \neg W_a p)$
$G_i \Diamond (p \wedge \neg W_i p)$	und	$G_i \Diamond (p \wedge \forall a \neg W_a p)$

- Inkonsistent ist hingegen:

(*) $G_i(p \wedge \neg W_i p)$

Mit Glauben legen wir uns fest, *nicht nicht zu wissen*.

[Warum ist (*) inkonsistent?]

- $\vdash \neg G_i(p \wedge \neg W_i p)$

Denn angenommen: $G_i(p \wedge \neg W_i p)$

dann [$\wedge$Dis]: $G_i p \wedge G_i \neg W_i p$

also [Df. „W“]: $G_i p \wedge G_i(\neg G_i p \vee \neg p)$

also [l-Wissen]: $G_i p \wedge G_i(p \supset \neg G_i p)$

also [EA]: $G_i(\neg G_i p)$

also [S4]: $G_i G_i p$

also [$\wedge$Dis]: $G_i \bot$ ↯

Zur Erinnerung: $G_i \bot$ widerspricht dem Axiom $G_i p \supset \neg G_i \neg p$

D.h. solange wir G-Konsistenz annehmen, gilt das Theorem.

- Als Verallgemeinerungen:

Inkonsistent: $\exists a \exists p\ (W_a p \wedge W_a(\neg \exists b W_b p))$

Konsistent: $\exists a \exists t\ W_a(\neg \exists b W_b t)$

Der erste Satz ist inkonsistent, da wir nicht von einer *bestimmten* Wahrheit, die wir als solche kennen, *de dicto* wissen können, dass sie niemand weiß.

Der zweite Satz ist konsistent, da wir *de re* von einer bestimmten Wahrheit wissen können, dass sie niemand weiß.

Nichtallwissenheit

- Das Nichtallwissenheitsaxiom besagt: $\forall a \exists t \neg W_a t$
- Daraus folgen modallogisch:

$$\Box \forall a \exists t \neg W_a t$$

$$\forall a \Box \exists t \neg W_a t$$

Wissbarkeitsthesen

• Wissbarkeitsthesen (‚knowability theses‘) können in verschiedenen modalen Formen auftreten:

(i) $\forall t \forall a \Diamond W_a t$

(ii) $\forall t \Diamond \forall a\, W_a t$

(iii) $\forall t \exists a \Diamond W_a t$

(iv) $\forall t \Diamond \exists a\, W_a t$

Dabei gelten prädikatenlogisch (i) $\vdash$ (iii), (ii)$\vdash$(iv) und modallogisch (iii) $\vdash$(iv).
(iv) ist also die schwächstmögliche Formulierung.

- $\forall t \Diamond \exists a\, W_a t \vdash \forall t \exists a\, W_a t$ (‘Modal Collapse Theorem’)
- $\forall t \exists a\, W_a t$ widerspricht dem *Theorem der Nicht-gewussten-Wahrheit*: $\vdash \exists t \forall a\, \neg W_a t$

[Dieses wird in Kap. 8 hergeleitet.]

Also [Modal Collapse Theorem]: $\vdash (\neg \forall t \Diamond \exists a\, W_a t)$

Also: Alle Wissbarkeitsthesen (i) – (iv) sind abzulehnen!

- Zu begründen ist noch das *Modal Collapse Theorem*.

[Es heißt *Modal Collapse Theorem*, da mittels $\forall t \Diamond \exists a\, W_a t$,
d.h. $p \supset \Diamond \exists a\, W_a p$, gezeigt wird, dass $\forall t \exists a\, W_a t$, d.h. $p \supset \exists a\, W_a p$ gilt, so dass, gegeben Veritativität ($\exists a\, W_a p \supset p$), also $p \equiv \exists a\, W_a p$, d.h. Kollaps der Modalität ‚von jemand gewusst werden‘. Zurückgehend auf J.J. MacIntosh.]

Modal Collapse Theorem

- Angenommen: (1) $\forall t \Diamond \exists a\, W_a t$

  Angenommen: (2) $p \wedge \neg \exists a\, W_a p$

also [Anwendung (1) auf (2)]: $\Diamond \exists a\, W_a(p \wedge \neg \exists b\, W_b p)$

also [$\wedge$Dis]: $\Diamond \exists a\, (W_a p \wedge W_a(\neg \exists b\, W_b p))$

aber [wie oben gezeigt]: $\neg \Diamond \exists a \exists p\, (W_a p \wedge W_a(\neg \exists b W_b p))$

also [$\neg$E]: $\neg(p \wedge \neg \exists a\, W_a p)$

also [AL]: $p \supset \exists a\, W_a p$

also [Df. „t“,$\forall$E]: $\forall t\, \exists a\, W_a t$

Modal Collapse Fitch Style

- Dass man Nichtwissen nicht wissen kann, wird oft ‚Fitchs Paradox‘ oder ‚Paradox des Wissenden‘ genannt. Chalmers nennt solche Wahrheiten (bezüglich des eigenen Nichtwissens) ‚Fitchian Truth‘, denen ein epistemologischer Sonderstatus zukommt.
- Fitch zeigt in seinem Aufsatz – mutmaßlich wegen eines Hinweis des Referees Church – dass Wissbarkeitsthesen problematisch sind wegen einer Form des modalen Kollapses: mögliches Wissen führt zu tatsächlichem Wissen.

- Das Argument entspricht dem oben gegebenen in der Variation „W_a“ statt „$\exists a W_a$“ als Modalität. Man erhält das Konditional: $\forall t \Diamond W_a t \supset \forall t\, W_a t$
- Der Schluss von Möglichkeit auf Wirklichkeit ist – offensichtlich – im Allgemeinen problematisch.
- Insbesondere erscheint es bizarr, dass aus der *Wissbarkeit* jeder Wahrheit folgen soll, dass jede Wahrheit (von einem bestimmten Subjekt) *gewusst wird*.
- Hinweis: Das Konditional $\forall t \Diamond W_a t \supset \forall t\, W_a t$ enthält prädikatenlogisch *nicht* das wesentlich stärkere und noch unhaltbarere: $\forall t(\Diamond W_a t \supset W_a t)$.

• Gegeben, dass ich nicht alles weiß, $\neg\forall t\, W_i t$, folgt per Kontraposition aus dem Fitch-Konditional: $\neg\forall t \Diamond W_i t$. Meine Wissensmöglichkeiten sind also beschränkt.

• Das gilt für alle Subjekte. Also: $\forall a \neg\forall t \Diamond W_a t$, also [$\forall/\exists$, $\Box/\Diamond$]: $\forall a \exists t \Box \neg W_a t$. Dies weist wieder auf das *Theorem der nicht-gewussten Wahrheit*, sagt aber natürlich weniger, da hier die notwendigerweise nicht-gewusste Wahrheit individuell sein kann.

• Gegeben das Axiom der Nichtallwissenheit, $\forall a \exists t \neg W_a t$, lässt sich ein Argument für *Fitchs Theorem* geben. [Wie?]

- $\vdash \forall a \exists t \neg W_a t$

also [∃B]: $\exists t \neg W_a t$

also [∃B]: $\neg W_a t$

also [„$W_i(t \wedge \neg W_i t)$" ist inkonsistent]: $\neg W_a(t \wedge \neg W_a t)$

also [□E]: $\Box \neg W_a(t \wedge \neg W_a t)$

also [□/◇]: $\neg \Diamond W_a(t \wedge \neg W_a t)$

also [∃E, „$t \wedge \neg W_a t$" ist wahr]: $\exists t \neg \Diamond W_a t$

also [⊃E]: $\exists t \neg W_a t \supset \exists t \neg \Diamond W_a t$

also [Kontraposition, ∀/∃]: $\forall t \Diamond W_a t \supset \forall t W_a t$

also [∀E]: $\forall a(\forall t \Diamond W_a t \supset \forall t W_a t)$

Nichtwissbare Wahrheiten

- Eine nichtwissbare Wahrheit kann natürlich nicht konkret angegeben werden. Obwohl gilt

$$W_i(\exists t \forall a \neg \Diamond W_a t)$$

Dabei handelt es sich um das generische Wissen ums Nichtwissen.

- Kenne ich eine Wahrheit t_1, die *a* nicht kennt, weiß ich, das *a* notwendigerweise „$t_1 \wedge \neg W_a t_1$“ nicht wissen kann, d.h.

$$W_i(\exists t(W_i t \wedge \neg \Diamond W_a t))$$

Nichtinstantiierbare Prädikatoren

• Nichtwissen kann sich aus Kennzeichnungen von Eigenschaften ergeben. Etwa:

- ist eine Idee, die niemand je haben wird
- ist ein Ereignis, das in keinem Geschichtsbuch auftritt
- ist eine Zahl, die noch nie gedacht wurde

  usw.

• Die Entitäten, auf die diese Eigenschaften zutreffen, können *nicht direkt* angegeben werden. Trotzdem kann für eine solche Eigenschaft *F* wahr sein: $\exists x\ F(x)$

- Dies ähnelt der ω-Inkonsistenz. „$\exists x\ F(x)$“ ist wahr und wissbar, aber für jedes x_1 etc. ist „$F(x_1)$“ nicht wissbar.
- Das betrifft den Unterschied zwischen

$\exists a W_a \exists x F(x)$ [etwas Allgemeines wissen]

und

$\exists x(F(x) \wedge \exists a W_a F(x))$ [von einem Objekt etwas wissen]

- Die nicht *instantiierbaren* Prädikatoren oben sind nicht instantiierbar, da ihre Definition auf Nichtwissen verweist. Von einer Instanz zu wissen, ähnelt dann: $W_i(p \wedge \neg W_i p)$.
- Es gilt i.d.R. $\exists x(F(x) \wedge \neg \Diamond \exists a W_a F(x))$

• Wir können auch von der Nichinstantiierbarkeit wissen. Sei

$$F^*(\ ) \stackrel{\text{def}}{=} F(\ ) \wedge \neg\Diamond\exists a W_a F(\ ))$$

dann erfüllt ein Objekt *u* diesen Prädikator, wenn es die nicht instantiierbare Eigenschaft *F* hat und dies *von ihm* nicht gewusst werden kann.

• Dies wiederum können wir allerdings wissen. Konsistent sind:

$$\exists a W_a F^*(u)$$

$$\exists a W_a(\exists x(F(x) \wedge \neg\Diamond\exists a W_a F(x)))$$

Fragen

- Eine Frage *f* ist beantwortbar, wenn es eine Antwort *gibt*: $\exists p(\text{beantwortet}(f,p)$. Nur solche Fragen sind sinnvoll.
- Ich habe eine Frage beantwortet, wenn ich die Antwort weiß: $W_i p \wedge \text{beantwortet}(f,p)$ [*de re*] oder sogar: $W_i(p \wedge \text{beantwortet}(f,p))$ [*de dicto*]
- Das Nichtallwissenheitsaxiom ist kompatibel mit der Existenz eines Oberschlaumeiers:

$$W_o(\text{beantwortet}(f,p) \equiv \Diamond\exists a W_a(\text{beantwortet}(f,p))$$

Nichtbeantwortbare Fragen

- Nichtbeantwortbare Fragen ähneln den Paradoxien bzw. den nicht instantiierbaren Prädikatoren.
- Beispiele:
    - Wirst Du diese Frage verneinen?
    - Was ist ein Beispiel für einen Satz, den Du nie sagst?

    usw.

- Wieder können wir generisch von unserer Unwissenheit wissen:

$$W_i(\exists f(\neg\exists p(W_i(\text{beantwortet}(f,p))))))$$

$$\exists f(W_i(\neg\exists p(W_i(\text{beantwortet}(f,p))))))$$

- Bei einzelnen Fragen f_1, f_2 … können wir auch von unserem Nichtwissen wissen:

$$W_i(\neg\exists p(W_i(\text{beantwortet}(f_1,p))))))$$

Allgemeines Unwissen

- Mit einer Argumentationsweise wie bei derjenigen zu nicht gewussten Wahrheiten, können wir argumentieren:

  (i) Für jedes epistemische Subjekt gibt es eine Frage, die es nicht beantworten kann.

  (ii) Also gibt es eine konjungierte (große) Frage, die kein epistemisches Subjekt beantworten kann.

- Die Plausibilität von (i) führt damit zu:

$$\exists f \neg \exists a \exists p \Diamond W_a(\text{beantwortet}(f,p))$$

- Wiederum können wir kein Beispiel angeben.