

# EPISTEMISCHE LOGIK

## Grundlagen

### Meta-Wissen

Manuel Bremer  
University of Düsseldorf, Germany  
[www.mbph.de](http://www.mbph.de)

# Meta-Wissen der Ersten vs. Dritten Person

- Meta-Wissen kann einmal vorliegen als Wissen einer Person (*ego*) über die epistemischen Zustände einer anderen Person (*alter*). In der Epistemischen Logik spielt dieses Wissen i.d.R. eine untergeordnete Rolle.
- Introspektionsprinzipien betreffen die Zugänglichkeit der eigenen Zustände für einen selbst. Diese Prinzipien ähneln den Iterationsprinzipien der Modallogik

$$\Box p \supset \Box \Box p \quad \text{oder} \quad \Diamond p \supset \Box \Diamond p$$

- Introspektionsprinzipien können allerdings – insbesondere im Kontext von Logischer Allwissenheit – höchst implausible Konsequenzen haben. Sie werden i.d.R. in der Epistemischen Logik angenommen. Das mag plausibler sein für automatische Wissensverarbeitung.
- Die Probleme im Kontext Logischer Allwissenheit können hier nicht auftreten, da diese nicht angenommen wird (s.o.).
- Iterationsprinzipien machen, obwohl *einige* epistemische Akte bewusst sind und alle von Reflexionsakten begleitet werden *können*, wenig Sinn in einer Akt-Lesart der epistemischen Zustände – wegen geistigen ‚Überlaufs‘.  
[Warum bzw. d.h.?)

- Würde man  $G_{ap} \supset G_a G_{ap}$  als das tatsächliche Vollziehen eines weiteren Glaubensaktes verstehen, müsste dieser ja auch wieder geglaubt werden:  $G_a G_{ap} \supset G_a G_a G_{ap} \dots$

So scheint der Geist an Glaubensakten überzulaufen.

- Hier zeigt sich (wieder), dass die Epistemische Logik nicht mit einer Philosophie des Geistes, welche sich um eine Theorie der Bewusstheit und der Reflexion kümmern muss, verwechselt werden darf. Beides leistet sie nicht!

- *Dispositional* machen die Iterationsprinzipien Sinn: Sie drücken Zugänglichkeit der eigenen epistemischen Zustände aus. Dies wird oft erkenntnistheoretisch verlangt (z.B. kann man die Kohärenz eigener Meinungen nur prüfen, wenn man Meinungen über sie bilden kann und sie einem selbst zugänglich sind.)

[So formuliert etwa Bonjour in seiner Kohärenztheorie des empirischen Wissens eine ‚doxastic presumption‘ der Zugänglichkeit.]

# Spezifisches vs. generisches Meta-Wissen

- Meta-Wissen kann Wissen über Wissen spezifischer Subjekte sein  $W_a G_b p$  oder es kann sich um generisches Meta-Wissen handeln, etwa  $\forall a W_a \forall b W_b (p \vee \neg p)$ .
- Generisches Meta-Wissen ist eine Komponente einer Theorie des kollektiven Wissens bzw. einer Theorie des geteilten (kollektiven) Wissens.

## Meta-Wissen der Dritten Person

- Wenn *alter* etwa nicht weiß, kann *ego* dies problemlos wissen:  $W_{a \neg} W_{bp}$ . Es besteht auch kein Widerspruch zwischen nicht geteiltem Wissen:  $W_{a \neg} W_{bp} \wedge W_{ap}$
- Es gilt ein Prinzip der Wissen Kooptation:

Theorem  $\vdash W_a W_{bp} \supset W_{ap}$

denn man kann nur wissen, dass *alter* etwas *weiß*, wenn man selbst *weiß*, dass *alter* mit seinem Glauben richtigliegt.

[Beweis?]

- $\vdash W_a W_b p \supset W_a p$

Angenommen:  $W_a W_b p$

also [Df. „ $W$ “]:  $W_a(G_b p \wedge p)$

also [ $\wedge$ Dis,  $\wedge$ B]:  $W_a p$

- also auch:

Theorem  $\vdash W_a W_b p \supset W_b p \wedge W_a p$

- und etwas generischer:

1-Theorem  $\vdash W_a \exists b W_b p \supset W_a p$

vorausgesetzt werden muss nur, dass man ( $\exists$ B) kennt.



# Geheimnisse

- Ähnlich wie  $W_a \neg W_b p \wedge W_a p$  gibt es auch Geheimnisse:

$$W_i(t \wedge \neg \exists b (b \neq i \wedge W_b t))$$

- Epistemische Subjekte lassen sich auch anhand ihrer Geheimnisse individuieren.

# Experten

- Spiegelbildlich zu meinem privilegierten Wissen meiner Geheimnisse oder der Ignoranz anderer, kann ich auch wissen, dass ein anderer ein Experte bezüglich einer Frage ist (d.h. Wissen hat), wo ich keins habe:

$$W_i(\exists a)((W_{ap} \vee W_{a\neg p}) \wedge \neg W_{ip} \wedge \neg W_{i\neg p})$$

- Nicht wissen kann ich hingegen:  $W_i(W_{ap} \wedge \neg W_{ip})$   
[Warum?]

- $\neg W_i(W_{ap} \wedge \neg W_{ip})$

denn [ $\wedge$ Dis]

$$W_i(W_{ap} \wedge \neg W_{ip}) \supset W_i W_{ap} \wedge W_i \neg W_{ip}$$

also aus dem eben ausgeführten Kooptationsprinzip und Veritativität:

$$W_i(W_{ap} \wedge \neg W_{ip}) \supset W_{ip} \wedge \neg W_{ip}$$

$$W_i(W_{ap} \wedge \neg W_{ip}) \supset \perp$$

$$\neg W_i(W_{ap} \wedge \neg W_{ip})$$

# Jemandes Folgerungen kennen

- Insofern wir die Prinzipien Epistemische Abtrennung und Konjunktivität kennen, können wir von den Annahmen Watsons und seinem Wissen über (materiale) Implikationen auch auf *seine* (d.h. die von ihm gezogenen Konsequenzen) schließen:

$$G_a G_b p \wedge G_a G_b (p \supset q) \supset G_a G_b q$$

- Schlüsse dieser Art dienen auch der Kritik anderer bezüglich ihrer Inkonsequenz: Wir sehen Konsequenzen, die sie (als rationale Subjekte) auch sehen sollten.

# Positive Introspektion

- Positive Introspektion entspricht dem S4-Axiom

$$\Box p \supset \Box \Box p$$

- Gegeben die Vielfalt der epistemischen Operatoren bietet sich eine Vielzahl von Prinzipien an, nicht nur für jeweils einen Operator, sondern auch für die Mischung von Operatoren.
- Welche Prinzipien man annimmt hängt des Weiteren vom zugrunde gelegten Wissensbegriff ab. Hier wird ja von einer schwachen Wissensdefinition ausgegangen.

- (1)  $W_{ap} \supset W_a W_{ap}$       [sogenannte ‚KK-These‘]
- (2)  $\ddot{U}_{ap} \supset \ddot{U}_a \ddot{U}_{ap}$
- (3)  $G_{ap} \supset G_a G_{ap}$
- (4)  $M_{ap} \supset M_a M_{ap}$
- (5)  $G_{ap} \supset \ddot{U}_a G_{ap}$
- (6)  $M_{ap} \supset \ddot{U}_a M_{ap}$
- (7)  $\ddot{U}_{ap} \supset W_a \ddot{U}_{ap}$
- (8)  $G_{ap} \supset W_a G_{ap}$
- (9)  $M_{ap} \supset W_a M_{ap}$

[Welche von diesen Prinzipien sollte man annehmen?  
Welche von diesen Prinzipien sind fundamental?]

- Angenommen unser Axiom  $\ddot{U}_{ap} \supset G_{ap}$  und das Theorem  $G_{ap} \supset M_{ap}$  sind (3) und (4) bloße Abschwächungen von (5) und (6).

- Angenommen die ‚doxastic presumption‘, sollte als Axiomenschema gelten

Axiom für einen epistemischen Zustand  $\varphi$

$$\varphi_{ap} \supset \ddot{U}_a \varphi_{ap}$$

Zur Erinnerung: es geht ‚nur‘ um dispositionale Zugänglichkeit.

- Dieses Axiom deckt die Fälle (2), (5), (6) ab. (3) und (4) sind dann Theoreme.

# Positive Introspektion und Wissensbegriff

- Gegeben den schwachen Wissensbegriff sind (1), (7), (8) und (9) ebenfalls Theoreme. [Warum?]



- $\vdash W_{ap} \supset W_a W_{ap}$

Angenommen:  $W_{ap}$

also [PosIntr]:  $\dot{U}_a W_{ap}$

also [per Axiom]:  $G_a W_{ap}$

also [Df. „W“]:  $W_a W_{ap}$

- entsprechend für (7), (8), (9).

- Ist dies nicht zu stark? Sagt nicht der schwache Wissensbegriff, dass man zufällig etwas wissen kann, d.h. ohne sich seines Wissens sicher zu sein? Für  $I_{ap}$  gilt sicher  $I_{ap} \supset I_a I_{ap}$ , da es sich um das Vorliegen von Information zu den vorliegenden Informationen handelt (gegeben Iteration von „I“ ist syntaktisch zulässig). [Sind (1) etc. zu stark?]

- Zur Erinnerung  $W_a p$  drückt aus, dass  $a$  den *dispositionalen* Glauben, dass  $p$ , hat und  $p$  vorliegt.
- Angenommen,  $a$  hat nur zufällig recht. Dann hat  $a$  zufällig Wissen (auch ohne dass zu erkennen). Hat  $a$  nun auch den Glauben,  $p$  zu glauben, dann glaubt  $a$  – im schwachen Sinne – auch  $p$  zu wissen. (Das wiederholt die Idee, dass wir, wenn wir etwas glauben – d.h. für mehr als bloß möglich halten – nie glauben können, es nicht zu wissen.) Also, da  $a$  ja *zufällig* weiß, dass  $p$ , weiß  $a$  *auch zufällig*, dass  $a$  weiß, dass  $p$ . Dieses Meta-Wissen muss aber indessen ebenso wenig erkennen wie sein ursprüngliches Wissen. Jemand kann – zumindest gegeben den schwachen Wissensbegriff – Wissen *haben*, ohne dass *ihm dies transparent ist*.

- Mit ‚nicht erkennen‘ ist hier gemeint: ein epistemisches Subjekt hat den Zustand des Wissens ist sich aber nicht *bewusst*, dass es sich um Wissen handelt. Ohne diese Möglichkeit, gäbe es auch nicht die Möglichkeit, überhaupt von zufälligem Wissen zu reden.
- Gegeben dispositionale epistemische Zustände und den schwachen Wissensbegriff kann man ‚Positive Introspektion‘ akzeptieren.
- Für einen starken Wissensbegriff, der eine Rechtfertigungsbedingung enthält, wird Positive Introspektion auch fragwürdig sein.

- Es sei also betont: da hier *nicht* eine Theorie der Bewusstheit und Reflexion (eine Philosophie des Geistes), sondern eine Theorie epistemischer Zustände als Dispositionen und ihrer logischen Zusammenhänge (eine Epistemische Logik) entwickelt wird, kann man hier stärkere Annahmen der Introspektion machen.
- Hilfreich kann sein, statt von ‚Introspektion‘ von ‚Zugänglichkeit‘ zu sprechen, insofern ‚Introspektion‘ die Idee der Bewusstheit nahelegt (insofern der Begriff ja gerade der Philosophie des Geistes entlehnt wurde).

# Quasi-Mooresche Sätze

- Wer den schwachen Wissensbegriff kennt, kann nicht einfach sagen:

(1) Ich glaube, dass  $p$ , aber ich weiß es nicht.

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

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

- Da Wissen ein externalistischer Begriff ist, wird mit (1\*) i.d.R. zu viel beansprucht. Woher weiß ich, nicht zu wissen?
- Mit (1\*\*) droht ein Quasi-Moorescher Widerspruch.  
[Warum?]

• Angenommen:  $G_i p \wedge \neg W_i p$

Angenommen:  $p$

dann [Df „W“]:  $W_i p \not\Leftarrow$

• Angenommen:  $G_i p \wedge G_i \neg W_i p$

also [Positive Introspektion]:  $G_i G_i p$

also [ $\wedge$ Dis]:  $G_i(G_i p \wedge p)$

also [Kenntnis Df „W“]:  $G_i W_i p$

also [ $\wedge$ Dis]:  $G_i(W_i p \wedge \neg W_i p) \not\Leftarrow$

also [ $\neg$ E]:  $\neg(G_i p \wedge G_i \neg W_i p)$

d.h.  $\vdash G_i p \supset \neg G_i \neg W_i p$

• Daraus folgt natürlich *nicht*:  $G_i p \supset G_i W_i p$  [Warum?]

- $\not\vdash (G_i p \supset \neg G_i \neg W_i p) \supset (G_i p \supset G_i W_i p)$

denn Glauben ist nicht maximal:  $\neg(G_i p \vee G_i \neg p)$

- Mit Positiver Introspektion erübrigen sich allerdings iterierte Wissensbehauptungen, es gilt ja:

$$\vdash W_i p \equiv W_i W_i p$$

Mit

- (1) Ich weiß, dass ich weiß, dass  $p$ .

sage ich also nicht mehr als mit

- (2) Ich weiß, dass  $p$ .

Deshalb kommen Sätze des Typs von (1) selten vor.

# Negative Introspektion

- Negative Introspektion entspricht dem S5-Axiom

$$\diamond p \supset \square \diamond p$$

- Ausgedrückt mit dem Wissensoperator:

$$\neg W_a p \supset W_a \neg W_a p$$

- Analog zur Positiven Introspektion gibt es verschiedene Variationen davon, etwa

$$\neg G_a p \supset G_a \neg G_a p$$



- Negative Introspektion ist völlig inakzeptabel, sowohl in einer Theorie der kognitiven Fähigkeiten als auch erkenntnistheoretisch. [Warum?]

- Kognitionstheoretisch: Jede Sprache enthält unendlich viele Sätze. Fast alle von diesen hat man noch nie erwogen. Hat man den Glauben, einen Satz  $\alpha$  nicht zu glauben, hat man  $\alpha$  gerade erwogen. Dies gilt für beliebige Sätze  $\alpha$ . Genau das sagt nun aber Negative Introspektion: man hat einen Meta-Glauben bezüglich beliebiger nicht-erwogener Sätze der Sprache, die man damit doch erwägt. Sie müssen alle irgendwie dennoch zugänglich sein und mittels entsprechender Zustände in das Rasonieren eingehen. Das überfordert mutmaßlich die Ressourcen endlicher epistemischer Subjekte.

- Erkenntnistheoretisch: Angenommen den schwachen Wissensbegriff, Positive Introspektion und Negative Introspektion ergibt sich mit geringem logischen Wissen eine Allwissenheit bezüglich beliebiger Sachverhalte!  
Das ist eine *reductio* für Negative Introspektion, insofern die anderen Annahmen akzeptiert wurden.

[Warum ergibt sich das?]

• Angenommen:  $\neg W_i p$

aber ich bin überzeugt oder glaube es:  $\ddot{U}_i p \vee G_i p$

also:  $G_i p$

also [Positive Introspektion]:  $W_i G_i p$

also [Df. „W“]:  $\neg p$

also [Negativer Introspektion]:  $W_i \neg W_i p$

also [Wissensaxiom, Kontraposition, EA]:  $W_i \neg (G_i p \wedge p)$

1-Postulat:  $W_i (\neg (G_i p \wedge p) \supset (G_i p \supset \neg p))$

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

also [EA]:  $W_i \neg p$

also [ $\supset E, \forall E$ ]:  $\vdash \forall a \forall p (\ddot{U}_i p \vee G_i p \supset (\neg W_a p \supset W_a \neg p))$

- $\vdash \forall a \forall p (\dot{U}_i p \vee G_i p \supset (\neg W_{ap} \supset W_{a\neg p}))$

ist zum einen epistemisch inakzeptabel, da es hieße, dass Wissen bezüglich dessen, was man für wahr hält, ‘maximal’ ist:  $W_{ap} \vee W_{a\neg p}$ , was jedoch offensichtlich nicht so ist.

- Zum anderen könnte man sich, indem man obiges Rasonieren bezüglich seines Glaubens durchläuft, *jede beliebige Frage beantworten*, wäre also allwissend – man weiß schon oder korrigiert sich einfach im Fall des Nichtwissens.

- Die erkenntnistheoretische *reductio* erstreckt sich nicht auf G-Varianten Negativer Introspektion, die indessen wegen des eher kognitionstheoretischen Einwandes ebenfalls abzulehnen sind.
- Negative Introspektion muss also abgelehnt werden.