Marius Thomann
marius.thomann@uni-koeln.de
Philosophisches Seminar
Universität zu Köln

XXII. Deutscher Kongress für Philosophie
Sektion: Logik
LMU München
15. 09. 2011

Dialetheismus, semantische Geschlossenheit und Konditionale

Ein Einwand gegen die Formulierung einer semantisch geschlossenen Sprache vor dem Hintergrund des sogenannten Dialetheismus zielt darauf ab, mittels eines verstärkten Lügnersatzes die Trivialität der Sprache nachzuweisen. Im vorliegenden Aufsatz wird eine entlang dieser Strategie vorgebrachte Kritik diskutiert. Im Zuge dieser Diskussion lassen sich Einschränkungen entwickeln, die das Konditional erfüllen muss, wenn die Sprache (i) den Grundannahmen des Dialetheismus genügen, (ii) nichttrivial und (iii) semantisch geschlossen sein soll. Als Ergebnis zeigt sich, dass Graham Priests Logik *LP*, die paradigmatische dialetheistische Logik, nicht durch den hier diskutierten Einwand gefährdet ist.

1 Einleitung

Das Unternehmen, eine semantisch geschlossene Sprache zu charakteriseren, d. h. eine Sprache, die wahrheitsgemäß ihre eigene Semantik charakterisiert, ist mit dem Problem der sogenannten *semantischen Paradoxien* konfrontiert. Am bekanntesten unter diesen ist vermutlich die *Lügnerparadoxie*: Wenn eine Sprache auf ihre eigenen Aussagen referieren kann und außerdem ein Wahrheitsprädikat enthält, das Tarskis (1936, S. 305 f.) Konvention W genügt, ist der *Lügnersatz*, der von sich selbst behauptet falsch zu sein, genau dann wahr, wenn er falsch ist. Unter der Annahme, alle Sätze seien wahr oder falsch, ergibt sich daraus, dass der Lügnersatz wahr und falsch zugleich ist. In seinem Aufsatz „The logic of paradox" hat Graham Priest eine ausgesprochen radikale Lösung für diese Problematik vorgeschlagen:

> Suppose we stop banging our heads against a brick wall trying to find a solution, and accept the paradoxes as brute facts. That is, some sentences are true (and true only), some false (and false only), and some both true and false. (1979, S. 220)

Die von Priest vertretene Position, manche Sätze seien wahr und falsch zugleich, wird als *Dialetheismus* bezeichnet. Ihre formale Modellierung verlangt nach einer *parakonsistenten Logik*, wobei eine Logik *L* parakonsistent ist, wenn *ex contradictione quodlibet*

($\varphi, \neg\varphi \Vdash \psi$) kein L-gültiges Schlussprinzip ist.[1] Priest bevorzugt die von ihm formulierte Logik *LP* (vgl. 1979; 2006). Die folgende Diskussion wird auf diese Logik beschränkt bleiben.

Vielen Versuchen, eine konsistente Lösung für die Lügnerparadoxie zu finden, d. h. ohne Annahme der dialetheistischen Position auszukommen, droht die sogenannte *Revenge-Problematik*. So lässt sich z. B. für solche Ansätze, denen zufolge der Lügnersatz wahrheitswertlos ist, ein verstärkter Lügnersatz formulieren, der von sich selbst behauptet, nicht wahr zu sein. Unter der Annahme, jede Aussage sei wahr oder nicht wahr, nimmt der so verstärkte Lügnersatz zwei miteinander unverträgliche semantische Werte zugleich an: Er ist wahr und zugleich nicht wahr. Die Schwierigkeiten wurden lediglich verschoben und nicht gelöst. Zahlreiche Einwände gegen den dialetheistische Position zielen darauf ab, auch diesen Lösungsansatz durch einen auf bestimmte Weise zu formulierenden verstärkten Lügnersatz zu blockieren – so auch die von Joachim Bromand (2002) vorgetragene Argumentation. Eine Untersuchung von Bromands Kritik und Priests (2006, S. 290) Erwiderung macht deutlich, dass der Verknüpfung „$\rightarrow$“, die für die Formulierung der semantischen Grundprinzipien von *LP* benötigt wird, eine zentrale Rolle in dieser Diskussion zukommt.

Im Folgenden wird Bromands Einwand und Priests Reaktion darauf kurz dargestellt (Abschnitt 3). Anschließend (Abschnitt 4) soll dieser Einwand vor dem Hintergrund einer bestimmten Ableitungsregel diskutiert werden, einer Art *reductio ad absurdum*. Wenn die Grundannahme, auf der Bromand sein Argument formuliert, aus einer dialetheistischen Position heraus nicht bestreitbar ist und somit die Kritik eine echte Bedrohung darstellt, ist die zu diskutierende Ableitungsregel zulässig in dem Sinne, dass die Korrektheit einer Beweistheorie für *LP* durch Hinzunahme der Regel nicht verloren gehen kann. Die Frage, ob diese Regel zulässig ist, wird sich auf die Frage nach einer adäquaten Modellierung der Verknüpfung „$\rightarrow$“ reduzieren, die danach untersucht wird (Abschnitt 5). Doch zunächst soll der formale Rahmen der Untersuchung vorgestellt werden.

2 Der formale Rahmen

Die in diesem Aufsatz diskutierte Objektsprache der Logik *LP* ist die der Prädikatenlogik erster Stufe mit Identität (vgl. Priest, 2006, Kap. 5), erweitert um eine Verknüpfung „$\rightarrow$“: Das Vokabular besteht also aus einer Menge $\mathcal{O} = \{\neg, \wedge, \rightarrow \forall\}$ von *Operatoren*, einer Menge $\mathcal{V}$ von *Variablen*, einer Menge $\mathcal{I}$ von *Individuenkonstanten* und einer Menge $\mathcal{P}$ von *Prädikatsymbolen*, wobei letztere ein zweistelliges Symbol „=“ enthält. Die Menge Φ_0 der *atomaren Formeln* und, darauf durch rekursive Klauseln aufbauend, die

[1] Nicht jede parakonsistente Logik beruht auf der dialetheistischen Grundannahme, manche Sätze seien wahr und falsch zugleich. Für eine feinkörnigere Definition und die Unterscheidung verschiedener Stufen von Parakonsistenz vgl. Carnielli & Marcos (2002) und Priest (2002).

Menge Φ der *wohlgeformten Formeln* sind in der üblichen Weise charakterisiert. Weitere Operatoren sind folgendermaßen als abkürzende Schreibweisen definiert:

$$\begin{aligned}
&\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)\\
&\varphi \leftrightarrow \psi := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)\\
&\exists v\varphi := \neg\forall v\neg\varphi
\end{aligned}$$

Um die semantische Geschlossenheit von *LP* zu erlauben, muss außerdem eine Möglichkeit bereitgestellt werden, mit Ausdrücken der Objektsprache auf die Elemente von Φ zu referieren und semantische Eigenschaften von diesen Elementen zu prädizieren. Aus Gründen der Einfachheit werde ich im Folgenden annehmen, dass es für jedes $\varphi \in \Phi$ einen Term $\ulcorner\varphi\urcorner \in \mathcal{I}$ gibt, dessen Denotat per definitionem die Aussage φ ist. Insbesondere werde ich fordern, dass es einen Term „λ" gibt, dessen Denotat der später (Abschnitt 3) zu formulierende verstärkte Lügnersatz ist. Darüber hinaus soll $\mathcal{I}$ zwei Konstanten „1" und „0" enthalten und $\mathcal{P}$ ein zweistelliges Prädikatsymbol „V". Die Bedeutung dieser zusätzlichen Symbole wird später (Abschnitt 2.2) durch bestimmte Forderungen eingeschränkt.

2.1 *Semantik*

Im Rahmen dieser Untersuchung sollen in allgemeiner Form mögliche Modellierungen der Verknüpfung „$\rightarrow$" diskutiert werden. Aus diesem Grund ist es nicht sinnvoll, einen konkreten Modellbegriff zu definieren. In Abhängigkeit von der semantischen Charakterisierung der Verknüpfung „$\rightarrow$" sind jeweils unterschiedliche Anforderungen an die Beschaffenheit eines Modells zu richten, etwa dass eine Menge möglicher Welten und eine Zugänglichkeitsrelation enthalten sind. Daher greife ich auf das Konzept eines *LP-Bewertungspunktes*[2] zurück, das folgendermaßen definiert ist:

Definition 1 (*LP*-Bewertungspunkt)
Ein *LP-Bewertungspunkt* $\mathfrak{P}$ ist ein Tupel von Parametern, so dass

(i) $\mathfrak{P}$ eine Menge $D \neq \emptyset$ enthält, die *Domäne* von $\mathfrak{P}$, und

(ii) jedem $X \in \mathcal{V} \cup \mathcal{I} \cup \mathcal{P}$ an $\mathfrak{P}$ gemäß den folgenden Bedingungen ein eindeutiger Wert, $X^{\mathfrak{P}}$, zugewiesen wird:
   (a) für alle $X \in \mathcal{V} \cup \mathcal{I}$: $X^{\mathfrak{P}} \in D$;
   (b) für alle n-stelligen $X \in \mathcal{P}$: $X^{\mathfrak{P}} = \langle E, A\rangle$, wobei $E \cup A = D^n$.

Für ein gegebenes $P \in \mathcal{P}$ werden E und A durch $E(P^{\mathfrak{P}})$ und $A(P^{\mathfrak{P}})$ bezeichnet. □

[2] Für eine gegebene Logik L ist ein *L-Bewertungspunkt* eine Folge von Parametern, die hinreichend ist, um den semantischen Wert jeder L-Aussage zu bestimmen (vgl. Belnap et al., 2001, S. 142).

Der semantische Wert einer LP-Aussage wird relativ zu einem LP-Bewertungspunkt $\mathfrak{P}$ bestimmt. Die folgenden Bedingungen (vgl. Priest, 2006, S. 77) legen fest, wann eine Aussage φ *wahr an* $\mathfrak{P}$ ist, in Zeichen $\mathfrak{P} \vDash^T \varphi$, und wann φ *falsch an* $\mathfrak{P}$ ist, in Zeichen $\mathfrak{P} \vDash^F \varphi$, wobei $\mathfrak{P}[v/d]$ für gegebenes $v \in \mathcal{V}$ und $d \in D$ ein Bewertungspunkt ist, der sich höchstens dadurch von $\mathfrak{P}$ unterscheidet, dass $v^{\mathfrak{P}} = d$:[3]

(At.T) $\mathfrak{P} \vDash^T P(t_1, \ldots, t_n)$ gdw. $\langle t_1^{\mathfrak{P}}, \ldots, t_n^{\mathfrak{P}} \rangle \in E(P^{\mathfrak{P}})$

(At.F) $\mathfrak{P} \vDash^F P(t_1, \ldots, t_n)$ gdw. $\langle t_1^{\mathfrak{P}}, \ldots, t_n^{\mathfrak{P}} \rangle \in A(P^{\mathfrak{P}})$

($\neg$.T) $\mathfrak{P} \vDash^T \neg\psi$ gdw. $\mathfrak{P} \vDash^F \psi$

($\neg$.F) $\mathfrak{P} \vDash^F \neg\psi$ gdw. $\mathfrak{P} \vDash^T \psi$

($\wedge$.T) $\mathfrak{P} \vDash^T \psi_1 \wedge \psi_2$ gdw. $\mathfrak{P} \vDash^T \psi_1$ und $\mathfrak{P} \vDash^T \psi_2$

($\wedge$.F) $\mathfrak{P} \vDash^F \psi_1 \wedge \psi_2$ gdw. $\mathfrak{P} \vDash^F \psi_1$ oder $\mathfrak{P} \vDash^F \psi_2$

($\forall$.T) $\mathfrak{P} \vDash^T \forall v \psi$ gdw. für alle $d \in D$: $\mathfrak{P}[v/d] \vDash^T \psi$

($\forall$.F) $\mathfrak{P} \vDash^F \forall v \psi$ gdw. für wenigstens ein $d \in D$: $\mathfrak{P}[v/d] \vDash^F \psi$

Diese Klauseln verallgemeinern die der klassischen Prädikatenlogik auf natürliche Weise. Es ist lediglich darauf zu achten, dass Wahrheit und Falschheit partiell unabhängig voneinander sind.[4] Die folgende Definition *logischer Folgerung* ist ebenfalls analog zu der entsprechenden Charakterisierung im Rahmen der klassischen Prädikatenlogik formuliert und fordert Wahrheitserhaltung von den Prämissen zur Konklusion an allen Bewertungspunkten:

Definition 2 (logische Folgerung)
Eine Aussage $\varphi \in \Phi$ ist genau dann eine *logische Folgerung* aus einer Menge $\Gamma \subseteq \Phi$, in Zeichen $\Gamma \Vdash \varphi$, wenn für alle LP-Bewertungspunkte $\mathfrak{P}$ gilt:

$$\text{wenn } \mathfrak{P} \vDash^T \psi \text{ für alle } \psi \in \Gamma, \text{ dann } \mathfrak{P} \vDash^T \varphi. \qquad \square$$

Die Verknüpfung „$\rightarrow$" soll vorerst unbestimmt bleiben. Die einzige Forderung, die ich nachfolgend direkt an die Charakterisierung richten werde, ist die Gültigkeit des *modus ponens*:

(MP) $\varphi, \varphi \rightarrow \psi \Vdash \psi$

[3] In (At.T) und (At.F) ist $P \in \mathcal{P}$ ein n-stelliges Prädikatsymbol und $\langle t_1, \ldots, t_n \rangle \in (\mathcal{V} \cup \mathcal{I})^n$ ein n-Tupel von Termen. Aus Gründen der Lesbarkeit habe ich auf entsprechende Angaben in (At.T) und (At.F) verzichtet.

[4] Wenn φ nicht wahr an $\mathfrak{P}$ ist, muss φ falsch sein an $\mathfrak{P}$. Daher sind Wahrheit und Falschheit bloß *partiell* unabhängig voneinander.

Die Bedeutung der Individuenkonstanten „1“ und „0“ sowie des Prädikatsymbols „V“ wird durch die folgenden Forderungen hinreichend eingeschränkt:

(W) $\quad V(\ulcorner\varphi\urcorner, 1) \leftrightarrow \varphi$

(F) $\quad V(\ulcorner\varphi\urcorner, 0) \leftrightarrow \neg\varphi$

(B1) $\quad V(\ulcorner\varphi\urcorner, 1) \vee V(\ulcorner\varphi\urcorner, 0)$

(B2) $\quad \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow (x = 1 \vee x = 0))$

(W) und (F) stellen sicher, dass „$V(\ulcorner\varphi\urcorner, 1)$“ und „$V(\ulcorner\varphi\urcorner, 0)$“ intuitiv als „‘$\varphi$’ ist wahr“ bzw. „‘$\varphi$’ ist falsch“ gelesen werden können. (B1) und (B2) stellen gemeinsam sicher, dass jede Aussage wahr oder falsch ist und es außerdem keine weiteren semantischen Werte gibt. Im Folgenden sollen ausschließlich *LP*-Bewertungspunkte diskutiert werden, die zulässig im Sinne dieser Forderungen sind, d. h. nur solche Bewertungspunkte, an denen (W), (F), (B1) und (B2) wahr sind.

3 Der verstärkte Lügnersatz

Der verstärkte Lügnersatz, der gegen den Dialetheismus in Stellung zu bringen ist, behauptet von sich selbst *ausschließlich falsch* zu sein. Eine mögliche Formulierung mit Hilfe des hier verwendeten Vokabulars lautet wie folgt (vgl. Bromand, 2002):

(λ) $\quad V(\lambda, 0) \wedge \forall x(V(\lambda, x) \rightarrow x = 0)$

Die Grundlage für Bromands Einwand bietet das folgende Aussageschema, dass die semantische Grundannahme des Dialetheismus zum Ausdruck zu bringen scheint, alle Aussagen seien wahr (und nur wahr) oder falsch (und nur falsch) oder wahr und falsch zugleich:

(†)
$$\begin{aligned}&(V(\ulcorner\varphi\urcorner, 1) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 1)) \vee \\ &(V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)) \vee \\ &\qquad (V(\ulcorner\varphi\urcorner, 1) \wedge V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow (x = 1 \vee x = 0)))\end{aligned}$$

Auf Basis von (†) ergibt sich für λ eine dreifache Fallunterscheidung. Das erste und das dritte Disjunkt beinhaltet $V(\lambda, 1)$. Mit (W) ergibt sich daraus $V(\lambda, 0) \wedge \forall x(V(\lambda, x) \rightarrow x = 0)$, also λ selbst. Aus dem zweiten Konjunkt von λ folgt dann über universelle Instanziierung und (MP) $1 = 0$. Aufgrund von (B1) ergibt sich $V(\ulcorner\varphi\urcorner, 1)$ und über (W) schließlich φ für alle $\varphi \in \Phi$ – Trivialität. Das zweite Disjunkt ist λ selbst, so dass sich über (W) $V(\lambda, 1)$ ergibt und damit wiederum, wie in den ersten beiden Fällen, $V(\ulcorner\varphi\urcorner, 1)$ und schließlich über (W) φ für alle $\varphi \in \Phi$.

Priest weist (†) zurück mit dem Hinweis darauf, dass die semantischen Grundannahmen durch (B1) und (B2) zum Ausdruck gebracht werden:

> If one does this [i. e. adopt (†) as an axiom], then the argument of course goes through; but there is no reason why one should endorse such an axiom. Bromand notes that dialetheists are committed to the claim that every sentence is true only, false only, or both true and false; and he suggests that this is what (**) [i. e. (†)] expresses. It does not; it is **Trichotomy** [i. e. (B1)[5]] that expresses this fact. True, it does not express the fact that the only candidates for the value of α [i. e. φ] are 1 and 0. But this is what (*) [i. e. (B2)] expresses. (Priest, 2006, S. 290)

Die zentrale Frage, an der die nicht-triviale Formulierung einer semantisch geschlossenen Sprache im Rahmen des Dialetheismus hängt, ist aber, ob es einen zulässigen *LP*-Bewertungspunkt gibt, an dem (†) nicht wahr ist. Diese Frage soll im Folgenden durch Diskussion einer Ableitungsregel diskutiert werden, deren Korrektheit aus der Annahme folgt, (†) sei wahr an allen *LP*-Bewertungspunkten. Unter der Korrektheit einer Ableitungsregel R verstehe ich hierbei die folgende Eigenschaft: Wenn S eine korrekte Beweistheorie ist, dann ist auch $S+R$ korrekt, wobei $S+R$ durch Hinzunahme von R zu S entsteht. Wenn sich also herausstellt, dass die im nächsten Abschnitt diskutierte Ableitungsregel nicht korrekt in diesem Sinne ist, kann auch (†) nicht in allen *LP*-Bewertungspunkten wahr sein.

4 Eine dialetheistisch akzeptable reductio?

Wird die folgende Ableitungsregel zu einer korrekten Beweistheorie für *LP* hinzugenommen, so lässt sich λ aus der leeren Menge ableiten und es ergibt sich schließlich die Ableitbarkeit aller *LP*-Aussagen:[6]

(RT) Wenn $V(\ulcorner\varphi\urcorner, 1) \vdash 1 = 0$, dann $\vdash V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$.

Um Trivialität zu vermeiden, ist (RT) als eine nicht zulässige Ableitungsregel zurückzuweisen. Eine notwendige Bedingung für die Zulässigkeit einer Ableitungsregel ist ihre Korrektheit in obigem Sinne. Als mögliche Strategie für die Zurückweisung von (RT) bietet sich also der Nachweis an, dass diese Regel nicht korrekt ist.[7] Vor dem Hintergrund dieser Überlegung lässt sich die Diskussion von (RT) auf die Frage

[5] Priests **Trichotomy** ist, in der hier verwendeten Notation, folgendermaßen formuliert (vgl. 2006, S. 287): $(V(\ulcorner\varphi\urcorner, 1) \wedge \neg V(\ulcorner\varphi\urcorner, 0)) \vee (V(\ulcorner\varphi\urcorner, 1) \wedge V(\ulcorner\varphi\urcorner, 0)) \vee (\neg V(\ulcorner\varphi\urcorner, 1) \wedge V(\ulcorner\varphi\urcorner, 0))$. (B1) ist logisch äquivalent zu **Trichotomy** in dem Sinne, dass diese beiden Aussagen wechselseitig auseinander folgen.

[6] Natürlich muss die Beweistheorie auch über ein gewisses Maß an deduktiver Stärke verfügen: Im Rahmen eines Kalküls des natürlichen Schließens beispielsweise sind die Konjunktions-Elimination, die universelle Instanziierung, der *modus ponens* für „$\rightarrow$“ und „$\leftrightarrow$“ und die voraussetzungslose Ableitbarkeit aller Instanzen des Schemas (W) erforderlich. Mit diesen Regeln ist zunächst $1 = 0$ aus $V(\lambda, 1)$ ableitbar und damit gemäß (RT) $V(\lambda, 0) \wedge \forall x(V(\lambda, x) \rightarrow x = 0)$, also ($\lambda$) selbst, aus der leeren Menge.

[7] Sollte (RT) zwar korrekt in obigem Sinne sein, aber aus anderen Gründen als nicht zulässig erachtet werden, ist die Beweistheorie unvollständig. Insbesondere wäre dann λ wahr an allen *LP*-Bewertungspunkten. Daher halte ich den Nachweis, dass (RT) nicht korrekt ist, für die einzig mögliche Strategie, der drohenden Trivialisierung zu entgehen.

nach einer adäquaten Modellierung der Verknüpfung „$\rightarrow$“ reduzieren, da der folgende Anfang eines entsprechenden Korrektheitsbeweises für (RT) ab einer bestimmten Stelle von der Charakterisierung des Operators „$\rightarrow$“ abhängt:

Angenommen, $V(\ulcorner\varphi\urcorner, 1) \vdash 1 = 0$ und es gibt einen zulässigen LP-Bewertungspunkt $\mathfrak{P}$, so dass $\mathfrak{P} \not\models^T V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$. Gemäß ($\wedge$.T) ergeben sich zwei Möglichkeiten:

(i) $\mathfrak{P} \not\models^T V(\ulcorner\varphi\urcorner, 0)$. Da $\mathfrak{P}$ zulässig ist, gilt $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 1) \vee V(\ulcorner\varphi\urcorner, 0)$ und somit $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 1)$.[8] Nach Annahme ergibt sich damit aber $\mathfrak{P} \models^T 1 = 0$ und somit die Wahrheit aller LP-Aussagen an $\mathfrak{P}$, insbesondere also $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$.

(ii) $\mathfrak{P} \not\models^T \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$. Gemäß ($\forall$.T) gilt $\mathfrak{P}[x/d] \not\models^T V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0$.

Die Zulässigkeit von (RT) und damit auch die Nicht-Trivialität von LP hängt also von der Modellierung des Konditionals ab.

Bevor im nächsten Abschnitt die Einschränkungen diskutiert werden, die sich aus diesem Resultat für das Konditional ergeben, sei hier noch der Nachweis erbracht, dass die Korrektheit von (RT) aus der Annahme folgt, (†) sei wahr an allen Bewertungspunkten:

Angenommen, (†) ist wahr an allen Bewertungspunkten und $V(\ulcorner\varphi\urcorner, 1) \vdash 1 = 0$ im Rahmen einer korrekten Beweistheorie. Dann gilt $V(\ulcorner\varphi\urcorner, 1) \Vdash 1 = 0$. Es sei $\mathfrak{P}$ ein beliebiger Bewertungspunkt. Gemäß (†) sind zwei Fälle möglich:

(i) $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 1)$. Nach Annahme gilt dann $\mathfrak{P} \models^T 1 = 0$. Mit (B1) resultiert $\mathfrak{P} \models^T V(\ulcorner\psi\urcorner, 1)$ und über (W) schließlich $\mathfrak{P} \models^T \psi$ für alle $\psi \in \Phi$, insbesondere also $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$.

(ii) $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$.

Da $\mathfrak{P}$ beliebig gewählt war, ist (RT) eine korrekte Ableitungsregel in obigem Sinne, falls (†) wahr ist an allen Bewertungspunkten.

Mit einer negativen Antwort auf die Frage, ob (RT) eine korrekte Ableitungsregel ist, lässt sich also nicht bloß diese Regel als unzulässig zurückweisen, sondern auch nachweisen, dass (†) nicht wahr an allen Bewertungspunkten ist und somit die von Bromand vorgebrachte Kritik am Dialetheismus ihr Ziel verfehlt.

[8] An dieser Stelle mag der Einwand erhoben werden, dass der disjunktive Syllogismus in parakonsistenten Systemen kein gültiges Schlussprinzip ist. Wenn die Metasprache selbst einer parakonsistenten Logik folgt, schließen sich $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0)$ und $\mathfrak{P} \not\models^T V(\ulcorner\varphi\urcorner, 0)$ nicht gegenseitig aus. Mit dieser Erwiderung lässt sich das hier skizzierte Argument allerdings nicht unterbrechen: Gelten nämlich sowohl $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0)$ als auch $\mathfrak{P} \not\models^T V(\ulcorner\varphi\urcorner, 0)$, so gilt insbesondere $\mathfrak{P} \models^T V(\ulcorner\varphi\urcorner, 0)$ und somit kann „$V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$“ nicht aufgrund der ausschließlichen Falschheit des ersten Konjunkts selbst ausschließlich falsch an $\mathfrak{P}$ sein.

5 Konditionale

Eine wahrheitsfunktionale Charakterisierung der Verknüpfung „→“ kommt für eine nicht-triviale semantisch geschlossene Sprache im Rahmen des hier skizzierten Formalismus' nicht in Frage, wie die folgende Argumentation zeigt. Für den Nachweis werden zwei Fälle betrachtet, $d = 1^{\mathfrak{P}}$ und $d \neq 1^{\mathfrak{P}}$, und es zeigt sich, dass in beiden Fällen nicht alle Instanzen der Aussageschemata (W), (F), (B1) und (B2) wahr an $\mathfrak{P}$ sind, was im Widerspruch zur Zulässigkeit von $\mathfrak{P}$ steht.

(i) Angenommmen, $d = 1^{\mathfrak{P}}$. Es soll also gelten $\mathfrak{P}[x/1^{\mathfrak{P}}] \nvDash^T V(\ulcorner\varphi\urcorner, x) \to x = 0$. Wenn $\mathfrak{P}[x/1^{\mathfrak{P}}] \vDash^T V(\ulcorner\varphi\urcorner, x)$, ergibt sich wie an früherer Stelle aufgrund von $V(\ulcorner\varphi\urcorner, 1) \Vdash 1 = 0$ die Wahrheit aller LP-Aussagen an $\mathfrak{P}[x/d]$, insbesondere also auch die von „$V(\ulcorner\varphi\urcorner, x) \to x = 0$“. Also $\mathfrak{P}[x/1^{\mathfrak{P}}] \nvDash^T V(\ulcorner\varphi\urcorner, x)$. Da außerdem $\mathfrak{P}[x/1^{\mathfrak{P}}] \nvDash^T x = 0$, ergibt sich für eine wahrheitsfunktionale Charakterisierung von „→“ die folgende Klausel, wobei $\mathfrak{Q}$ ein beliebiger zulässiger Bewertungspunkt sei:

(wf) wenn $\mathfrak{Q} \nvDash^T \psi_1$ und $\mathfrak{Q} \nvDash^T \psi_2$, dann $\mathfrak{Q} \nvDash^T \psi_1 \to \psi_2$.

Nun sei $d' \in D$, so dass $d' \neq 1^{\mathfrak{P}}$ und $d' \neq 0^{\mathfrak{P}}$. Dann gilt $\mathfrak{P}[x/d'] \nvDash^T x = 1 \vee x = 0$. Da $\mathfrak{P}$ zulässig ist, ergibt sich mit (B2) $\mathfrak{P}[x/d'] \vDash^T V(\ulcorner\varphi\urcorner, x) \to (x = 1 \vee x = 0)$. Somit gilt aber auch $\mathfrak{P}[x/d'] \nvDash^T V(\ulcorner\varphi\urcorner, x)$ und schließlich, gemäß (wf) und (∀.T) $\mathfrak{P} \nvDash^T \forall x(V(\ulcorner\varphi\urcorner, x) \to (x = 1 \vee x = 0))$, was im Widerspruch zur Zulässigkeit von $\mathfrak{P}$ steht.

(ii) Angenommen, $d \neq 1^{\mathfrak{P}}$. Dann ist der semantische Wert von „$x = 1 \vee x = 0$“ an $\mathfrak{P}[x/d]$ mit dem von „$x = 0$“ identisch und somit, da eine wahrheitsfunktionale Charakterisierung für „→“ angenommen wird, auch der Wert von „$V(\ulcorner\varphi\urcorner, x) \to (x = 1 \vee x = 0)$“ mit dem von „$V(\ulcorner\varphi\urcorner, x) \to x = 0$“. Also gilt $\mathfrak{P}[x/d] \nvDash^T V(\ulcorner\varphi\urcorner, x) \to (x = 1 \vee x = 0)$ und schließlich $\mathfrak{P} \nvDash^T \forall x(V(\ulcorner\varphi\urcorner, x) \to (x = 1 \vee x = 0))$, was im Widerspruch zur Zulässigkeit von $\mathfrak{P}$ steht.

Unter einer wahrheitsfunktionalen Charakterisierung von „→“ ist also entweder (RT) eine zulässige Ableitungsregel und LP somit trivial oder der Formalismus nicht semantisch geschlossen, da wenigstens eines der semantischen Grundprinzipien nicht wahr an allen Bewertungspunkten ist.

Für eine allgemeiner formulierte Einschränkung seien die folgenden Aussage- und Regelschemata betrachtet:[9]

(A1) $\varphi \to \varphi$

(A2) $(\varphi \wedge \psi) \to \psi$

[9] Diese Schemata charakterisieren das negationsfreie Fragment der sogenannten *basic affixing relevant logic*, *B* (vgl. Priest, 2008, S. 192 f.).

(A3) $((\varphi \to \chi) \land (\psi \to \chi)) \to ((\varphi \lor \psi) \to \chi)$

(A4) $((\varphi \to \psi) \land (\varphi \to \chi)) \to (\varphi \to (\psi \land \chi))$

(A5) $(\varphi \land (\psi \lor \chi)) \to ((\varphi \land \psi) \lor (\varphi \land \chi))$

(R1) $\varphi, \varphi \to \psi \vdash \psi$

(R2) $\varphi \to \psi \vdash (\chi \to \varphi) \to (\chi \to \psi)$

(R3) $\varphi \to \psi \vdash (\psi \to \chi) \to (\varphi \to \chi)$

(R4) $\varphi, \psi \vdash \varphi \land \psi$

Mit Hilfe dieser Schemata lässt sich $\varphi \to \chi$ aus $\varphi \to (\psi \lor \chi)$ und $(\varphi \land \psi) \to \chi$ ableiten:

1.	$\varphi \to (\psi \lor \chi)$	Ann.
2.	$(\varphi \land \psi) \to \chi$	Ann.
3.	$(\varphi \land \chi) \to \chi$	A2
4.	$(((\varphi \land \psi) \to \chi) \land ((\varphi \land \chi) \to \chi)) \to (((\varphi \land \psi) \lor (\varphi \land \chi)) \to \chi)$	A3
5.	$((\varphi \land \psi) \to \chi) \land ((\varphi \land \chi) \to \chi)$	R4: 2, 3
6.	$((\varphi \land \psi) \lor (\varphi \land \chi)) \to \chi$	R1: 4, 5
7.	$\varphi \to \varphi$	A1
8.	$((\varphi \to \varphi) \land (\varphi \to (\psi \lor \chi))) \to (\varphi \to (\varphi \land (\psi \lor \chi)))$	A4
9.	$(\varphi \to \varphi) \land (\varphi \to (\psi \lor \chi))$	R4: 1, 7
10.	$\varphi \to (\varphi \land (\psi \lor \chi))$	R1: 8, 9
11.	$(\varphi \land (\psi \lor \chi)) \to ((\varphi \land \psi) \lor (\varphi \land \chi))$	A5
12.	$((\varphi \land (\psi \lor \chi)) \to ((\varphi \land \psi) \lor (\varphi \land \chi))) \to ((\varphi \land (\psi \lor \chi)) \to \chi)$	R2: 6
13.	$(\varphi \land (\psi \lor \chi)) \to \chi$	R1: 11, 12
14.	$((\varphi \land (\psi \lor \chi)) \to \chi) \to (\varphi \to \chi)$	R3: 10
15.	$\varphi \to \chi$	R1: 13, 14

Nun sei die folgende Instanz dieser Ableitbarkeitsbeziehung betrachtet:

$$V(\ulcorner\varphi\urcorner, x) \to (x = 1 \lor x = 0), (V(\ulcorner\varphi\urcorner, x) \land x = 1) \to x = 0 \vdash \\ V(\ulcorner\varphi\urcorner, x) \to x = 0$$

Auf der Grundlage von $V(\ulcorner\varphi\urcorner, 1) \vdash 1 = 0$ ergibt sich $V(\ulcorner\varphi\urcorner, x) \land x = 1 \vdash x = 0$. Wenn also von $V(\ulcorner\varphi\urcorner, x) \land x = 1 \vdash x = 0$ zu $\vdash (V(\ulcorner\varphi\urcorner, x) \land x = 1) \to x = 0$ übergegangen werden darf, muss für die Nicht-Trivialität von *LP* wenigstens eines der obigen Schemata widerlegbar sein.

Priest (2006, S. 88) schlägt eine modale Charakterisierung für „→“ vor, unter der sich ein Gegenbeispiel für die Korrektheit von (RT) formulieren lässt. Ein Modell ist ein 5-Tupel $\langle W, R, w_0, D, v \rangle$, wobei $W \neq \emptyset$ eine Menge möglicher Welten ist, $R \subseteq W^2$

eine Zugänglichkeitsrelation, $w_0 \in W$ die aktuale Welt, $D \neq \emptyset$ die Domäne und v eine auf $W \times (\mathcal{I} \cup \mathcal{P})$ definierte Interpretationsfunktion. Ein Bewertungspunkt ist ein Tripel $\langle \mathfrak{M}, w, s \rangle$, wobei $\mathfrak{M}$ ein Modell ist, w eine mögliche Welt in $\mathfrak{M}$ und $s : \mathcal{V} \longrightarrow D$ eine Variablenbelegung. Die Bedingungen für Wahrheit und Falschheit von Aussagen der Form $\psi_1 \rightarrow \psi_2$ an einem solchen Bewertungspunkt sind folgendermaßen formuliert:

($\rightarrow$.T) $\mathfrak{M}, w, s \vDash^T \psi_1 \rightarrow \psi_2$ gdw. für alle $w' \in W$ mit wRw': wenn $\mathfrak{M}, w', s \vDash^T \psi_1$, dann $\mathfrak{M}, w', s \vDash^T \psi_2$

($\rightarrow$.F) $\mathfrak{M}, w, s \vDash^F \psi_1 \rightarrow \psi_2$ gdw. es ein $w' \in W$ mit wRw' gibt, so dass $\mathfrak{M}, w', s \vDash^T \psi_1$ und $\mathfrak{M}, w', s \vDash^F \psi_2$

Logische Folgerung ist definiert als Wahrheitserhaltung an allen Bewertungspunkten $\langle \mathfrak{M}, w_0, s \rangle$, wobei w_0 die aktuale Welt in $\mathfrak{M}$ ist. Diese Einschränkung und die obige Charakterisierung erlaubt die Formulierung des folgenden Gegenbeispiels für die Korrektheit von (RT):

$$\begin{aligned}
&W = \{w_0, w_1\},\\
&R = \{\langle w_0, w_0 \rangle, \langle w_0, w_1 \rangle\},\\
&\{\mathsf{t}, \mathsf{f}\} \subseteq D,\\
&v_{w_0}(1) = v_{w_1}(1) = \mathsf{t},\ v_{w_0}(0) = v_{w_1}(0) = \mathsf{f},\\
&\langle \varphi, \mathsf{f} \rangle \in v_{w_0}(V),\ \langle \varphi, \mathsf{t} \rangle \notin v_{w_0}(V),\\
&\langle \varphi, \mathsf{t} \rangle \in v_{w_1}(V),\ \langle \varphi, \mathsf{f} \rangle \notin v_{w_1}(V),\\
&s(x) = \mathsf{t}.
\end{aligned}$$

Es gelten $\mathfrak{M}, w_1, s \vDash^T V(\ulcorner \varphi \urcorner, x)$ und $\mathfrak{M}, w_1, s \nvDash^T x = 0$. Gemäß der oben angeführten Charakterisierung des Operators „$\rightarrow$“ resultiert $\mathfrak{M}, w_0, s \nvDash^T V(\ulcorner \varphi \urcorner, x) \rightarrow x = 0$. Mit dieser Konstruktion verträglich ist die Wahrheit von (W), (F), (B1) und (B2) an $\langle \mathfrak{M}, w_0, s \rangle$ sowie $V(\ulcorner \varphi \urcorner, 1) \Vdash 1 = 0$.

Da logische Folgerung nur an der jeweils aktualen Welt eines Modells wahrheitserhaltend sein muss, ist es möglich, trotz $V(\ulcorner \varphi \urcorner, 1) \Vdash 1 = 0$ eine von w_0 aus zugängliche mögliche Welt w' anzunehmen, nämlich w_1, so dass $\mathfrak{M}, w', s \vDash^T V(\ulcorner \varphi \urcorner, x)$ und $\mathfrak{M}, w', s \nvDash^T x = 0$. Es gibt also unter der von Priest vorgeschlagenen Modellierung von „$\rightarrow$“ einen zulässigen Bewertungspunkt, an dem „$V(\ulcorner \varphi \urcorner, 0) \wedge \forall x (V(\ulcorner \varphi \urcorner, x) \rightarrow x = 0)$“ falsch ist obwohl $V(\ulcorner \varphi \urcorner, 1) \Vdash 1 = 0$.

6 Zusammenfassung

Die Modellierung einer semantisch geschlossenen Sprache im Rahmen des hier skizzierten Systems wird nicht durch den verstärkten Lügnersatz λ trivialisiert. Bromands Schema (†) würde die Ableitung von „$1 = 0$“ aus der leeren Menge erlauben, ist allerdings selbst widerlegbar, d. h. nicht an allen zulässigen Bewertungspunkten wahr.

Wäre (†) nicht widerlegbar, so wäre (RT) eine korrekte Ableitungsregel im oben angeführten Sinne. Zwar ergeben sich aus der Diskussion von (RT) bestimmte Einschränkungen für die Charakterisierung der Verknüpfung „→“, doch formuliert Priest selbst Wahrheits- und Falschheitsbedingungen für „→“, unter denen sich nicht die Wahrheit von „$V(\ulcorner\varphi\urcorner, 0) \wedge \forall x(V(\ulcorner\varphi\urcorner, x) \rightarrow x = 0)$“ an allen zulässigen Bewertungspunkten aus $V(\ulcorner\varphi\urcorner, 1) \Vdash 1 = 0$ ergibt.

Was bedeutet das nun für die dialetheistische Position und die Bewertung von λ? Als Dialetheist lässt sich wahrheitsgemäß sagen, dass λ falsch ist. Und es lässt sich nicht sagen, λ sei wahr. Tatsächlich sagt der Dialetheist, λ sei falsch und nicht wahr: $V(\lambda, 0) \wedge \neg V(\lambda, 1)$. Und es gibt keinerlei Verpflichtung, außerdem auch die Wahrheit von λ zu behaupten. Die ausschließliche Falschheit einer Aussage wird im Rahmen des hier skizzierten Formalismus’ durch „$V(\ulcorner\varphi\urcorner, 0) \wedge \neg V(\ulcorner\varphi\urcorner, 1)$“ ausgedrückt.

> A dialetheist can express the claim that something, α, is not true—in those very words, $\neg T\langle\alpha\rangle$ [i. e. $\neg V(\ulcorner\alpha\urcorner, 1)$]. (Priest, 2006, S. 291)

Aus Perspektive einer klassisch konstruierten Modelltheorie lassen sich natürlich drei semantische Werte charakterisieren, die paarweise miteinander unverträglich sind: ausschließlich wahr, ausschließlich falsch und wahr und falsch zugleich. Und die in Abschnitt 1 zitierte Passage aus Priests „The logic of paradox“ legt eine derartige Konstruktion nahe. Die dialetheistische Grundannahme lässt sich aber auch ohne die Annahme inkompatibler semantischer Werte formulieren, nämlich indem lediglich gefordert wird, dass die zwei semantischen Werte, Wahrheit und Falschheit, einander nicht ausschließen. Im Sinne einer einschließenden Disjunktion sind Aussagen wahr oder falsch, und λ ist falsch. Punkt!

Es bleibt ein gewisses Unbehagen angesichts der Tatsache, dass im Rahmen des Dialetheismus’ die Wahrheit einer Aussage nie aus logischen Gründen ausgeschlossen werden kann. In der klassischen Logik gibt es keinen Bewertungspunkt, an dem alle Aussagen wahr sind. Somit kann φ bei Wahrheit von „$V(\ulcorner\varphi\urcorner, 1) \rightarrow 1 = 0$“ aus logischen Gründen nicht wahr sein. Der Dialetheismus hingegen lässt triviale Bewertungspunkte zu, an denen alle Aussagen wahr sind. Sie sollten sich natürlich aus pragmatischen Gründen verbieten, aber damit ist bei Behauptung von „$V(\ulcorner\varphi\urcorner, 1) \rightarrow 1 = 0$“ die Wahrheit von φ auch nur aus pragmatischen Gründen ausgeschlossen.

Literatur

Belnap, Nuel, Perloff, Michael & Xu, Ming 2001: *Facing the Future*, Oxford University Press, Oxford.

Bromand, Joachim 2002: „Why paraconsistent logic can only tell half the truth“, *Mind*, **111**:741–749.

Carnielli, Walter A. & Marcos, João 2002: „A taxonomy of C-systems“, in: Walter A. Carnielli, Marcelo E. Coniglio & Itala M. Loffredo D'Ottaviano (Hrsg.), *Paraconsistency – the Logical Way to the Inconsistent*, *Lecture Notes in Pure and Applied Mathematics*, Bd. 228, Marcel Dekker, New York.

Priest, Graham 1979: „The logic of paradox“, *Journal of Philosophical Logic*, **8**:219–241.

—— 2002: „Paraconsistent logic“, in: D. Gabbay & F. Guenthner (Hrsg.), *Handbook of Philosophical Logic*, Bd. 6, 2. Aufl., S. 287–393, Kluwer Academic Publishers, Dordrecht.

—— 2006: *In Contradiction*, 2. Aufl., Oxford University Press, Oxford.

—— 2008: *An Introduction to Non-Classical Logic*, 2. Aufl., Cambridge University Press, Cambridge.

Tarski, Alfred 1936: „Der Wahrheitsbegriff in den formalisierten Sprachen“, *Studia Philosophica*, **1**:261–405.