

FORMALLOGISCHE ANALYSE DES OPERANTEN KONDITIONIERENS

Peter Klimczak¹, Günther Wirsching²

¹Brandenburgische Technische Universität Cottbus-Senftenberg

²Katholische Universität Eichstätt-Ingolstadt
peter.klimczak@b-tu.de

Einleitung

Eine der großen Strömungen der amerikanischen Psychologie Anfang des zwanzigsten Jahrhunderts war der Behaviorismus, und dieser war zum großen Teil eine Stimulus-Response-Psychologie, die stark von der Suche nach Pawlow'schen Reflexen unterschiedlicher Art geprägt war. Burrhus F. Skinner, anfangs ebenfalls von Pawlows Untersuchungen begeistert, stellte aber bei Forschungen mit Ratten im Rahmen seiner Dissertation [1] fest, dass erwartete oder intendierte Folgen einer Handlung diese in vielen Fällen weitaus stärker beeinflussen als vor der Handlung bestehende Stimuli. Davon motiviert, wurde er zum Gegner der Stimulus-Response-Psychologie und begann, eine (funktionale) Theorie vom Lernen durch operantes Konditionieren zu entwickeln [2].

Aufbauend auf einer formallogischen Repräsentation lässt sich Skinners operantes Konditionieren als die Deduktion von Intentionen und Präskriptionen beschreiben. Zwar hat die empirische Psychologie eindringlich gezeigt, dass formallogisches Schließen nicht dem ‚natürlichen‘ Schließen von Menschen entspricht,¹ allerdings geht es uns im Folgenden nicht um menschliches Lernen und damit auch nicht um menschliches Schließen. Unser Ziel ist es operantes Konditionieren als mögliche Lernform technischer kognitiver Agenten beschreibbar zu machen.²

1. Formallogische Grundlagen

In der Philosophie wird zwar zurecht zwischen Sollen (Präskriptionen) und Wollen (Intentionen) unterschieden, weil anzunehmen ist, dass zwei nicht vereinbare, also widersprüchliche Sachverhalte nicht gesollt, dafür aber gewollt werden können [6]; doch gerade deshalb ist es im hier behandelten Kontext sinnvoll, Intentionen innerhalb der deontischen Modallogik,³ und damit innerhalb eines Kalküls zu modellieren, in welchem in der Philosophie nur Präskriptionen symbolisiert werden. Wenn nämlich Widersprüchliches, ohne dass ein formallogischer Widerspruch ableitbar ist, zugleich gewollt werden kann, so sind die deduzierten Wollenssätze ohne Wert: Ihnen kann (zeitgleich) unmöglich entsprochen werden. Da im Kontext von technischen kognitiven Systemen auch Gewolltes die Funktion von Verhaltensanweisungen hat

¹Vgl. hierzu für das konditionale Schließen überblicksmäßig [3], für das kausale Schließen als Sonderform des konditionalen Schließens [4]. Einen guten Überblick über das Verhältnis von Logik und Psychologie bietet [5].

²Auch wenn die von der psychologischen Forschung entwickelten Theorien und Modelle überzeugen, indem sie die Diskrepanz zwischen formallogischem und menschlichem Schließen erklären und letzteres sogar besser voraussagbar machen [5], sind sie für eine Verwendung im Kontext technischer kognitiver Agenten nur beschränkt geeignet. Zumeist basieren diese nicht auf Kalkülen, was aber die Voraussetzung für einen Einsatz in technischen Systemen darstellt. Gegenüber der Default-Logik oder probabilistischer Modelle hat die hier verwendete klassische Logik wiederum den Vorteil, dass sie recht leicht um deontische und ontische Modaloperatoren erweiterbar ist.

³Vgl. allgemein zur deontischen Modallogik [7] (kürzer) oder [8] (ausführlicher).

[9, 10], ist es nur folgerichtig, Intentionen wie Präskriptionen zu behandeln und mittels Sollsätzen abzubilden. Für die Darstellung von Sollsätzen verwendet die deontische Modallogik den **O**-Operator, mit dessen Hilfe Sachverhalte als gesollt und in unserem Fall nun auch als gewollt gesetzt werden können: $\mathbf{O}(p)$.⁴

Da in der deontischen Modallogik weder gilt, dass aus dem Gesolltsein von p das Sein von p folgt, noch gilt, dass aus dem Sein von p , das Gesolltsein von p folgt,⁵ sind keine Schlüsse aus Deskriptionen (Tatsachen) auf Präskriptionen und Intentionen möglich und umgekehrt. Allein aus Deskriptionen ergeben sich weder Präskriptionen noch Intentionen. Die Deduktion von (neuen) Sollsätzen ist lediglich unter Zuhilfenahme von Bedingungs- bzw. Folgerungsrelationen möglich⁶ — allerdings nur unter Hinzunahme bereits bestehender Präskriptionen und Intentionen. Entgegen linguistischen und argumentationstheoretischen Annahmen [12, 13] kann im Falle eines *Normenschlusses* nicht vom Sollen der Folge auf das Sollen der Ursache, also vom Sollen des Nachsatzes einer Subjunktion auf das Sollen des Vordersatzes, $\mathbf{O}(q) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(p)$, geschlossen werden. Das würde unweigerlich zu unmittelbaren und mittelbaren Widersprüchen und damit Fehlschlüssen schon in den einfachsten Kontexten führen.⁷ Widerspruchsfrei ist die Ableitung von mittelbaren Sollens- oder Wollensaussagen nur in umgekehrter Richtung möglich, vom Sollen/Wollen des Vordersatzes auf das Sollen/Wollen des Nachsatzes (deontischer modus ponens) oder entsprechend transponiert vom Sollen/Wollen der Negation des Nachsatzes auf das Sollen/Wollen der Negation des Vordersatzes (deontischer modus tollens):⁸

$$\mathbf{O}(p) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(q) \quad (1)$$

$$\mathbf{O}(\neg q) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(\neg p) \quad (2)$$

2. Operantes Konditionieren als Normenschluss

Aus psychologischer Sicht stellt Lernen einen Prozess dar, der als Ergebnis von Erfahrungen relativ langfristige Änderungen im Verhaltenspotential zur Folge hat [14]. Mittels der soeben beschriebenen Schlussverfahren lässt sich ein Modell entwerfen, das diese Definition erfüllt und damit Lernen für technische kognitive Agenten verwendbar macht. Dass ein operantes Konditionieren dabei Kognition nicht ausschließt, zeigen nicht zuletzt kognitive Ansätze in der psy-

⁴Während also **O** als „Es ist gesollt“ gelesen werden kann, steht p auch weiterhin für eine beliebige Aussage von Tatsachen, („Es gilt p “), sodass $\mathbf{O}(p)$ als „Es ist gesollt, dass p gilt“ zu lesen ist.

⁵Ganz im Gegensatz zur ontischen Modallogik, in der gilt, dass aus dem Notwendigsein von p das Sein von p folgt: $\mathbf{N}(p) \rightarrow p$.

⁶Zum Verhältnis von Bedingungs- und Folgerungsrelationen vgl. [11].

⁷Angenommen sei z. B. folgende (der Einfachheit aussagenlogisch repräsentierte) Situation, wobei „Kuchen-zu-essen-Bekommen“ als K , „Großmutter-Besuchen“ als G und „Tante-Besuchen“ als T symbolisiert seien: (α) Es besteht der Wunsch, Kuchen zu essen: $\mathbf{O}(K)$. (β) Man bekommt Kuchen zu essen, wenn man Großmutter besuchen geht, sodass eine Folgerungsrelation mit K als Folge und G als Ursache besteht: $\mathbf{N}(G \rightarrow K)$. (γ) Man bekommt auch Kuchen zu essen, wenn man die Tante besuchen geht: $\mathbf{N}(T \rightarrow K)$. (δ) Man kann (warum auch immer) nicht sowohl Tante als auch Großmutter besuchen gehen: $\neg\mathbf{M}(T \wedge G)$. Mittels der fraglichen Schlussregel $\mathbf{O}(q) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(p)$ könnte man aufgrund von α und β auf $\mathbf{O}(G)$ und aufgrund von α und γ auf $\mathbf{O}(T)$ schließen, was zusammen $\mathbf{O}(T \wedge G)$ entspricht. Aufgrund δ und der als gültig zu erachtenden Schlussregel (vgl. Anmerkung unten) $\neg\mathbf{M}(p) \rightarrow \neg\mathbf{O}(q)$, gilt allerdings auch $\neg\mathbf{O}(T \wedge G)$, was jedoch $\mathbf{O}(T \wedge G)$ widerspricht. Vgl. hierzu ausführlich [11].

⁸Neben dieser (letztlich) einzigen Möglichkeit, neue, also nicht bereits gesetzte oder gegebene, Präskriptionen oder Intentionen abzuleiten, gibt es noch zwei (lediglich funktionale – beide sind in komplexeren Kontexten für widerspruchsfreie Schlussverfahren nach deontischem modus ponens und deontischem modus tollens unverzichtbar) Verfahren: zum einen der Schluss auf das Nicht-Gesollt- oder Nicht-Gewollt-Sein von etwas, was nicht möglich ist, $\neg\mathbf{M}(p) \rightarrow \neg\mathbf{O}(p)$, zum anderen der Schluss auf das Gesollt- oder Gewollt-Sein von etwas, was notwendig ist, $\mathbf{N}(p) \rightarrow \mathbf{O}(p)$. Vgl. hierzu [10].

chologischen Forschung [15], die auch bei dieser Form des Lernens von kausalen Strukturen und logischen Schlussfolgerungen ausgehen.⁹

Entscheidend ist beim operanten Konditionieren, dass die Änderung des Verhaltenspotenzials operant bzw. instrumentell erfolgt: ein bestimmtes Verhalten bzw. eine bestimmte Reaktion des Lernenden wird seitens des Lehrenden belohnt oder bestraft mit der Folge, dass die Ausführung bzw. das Unterlassen dieses Verhaltens verstärkt, sprich erlernt, wird.¹⁰ Bestrafung und Belohnung bzw. bestrafende und belohnende Lehrendenreaktion können als zweistellige Prädikate formalisiert werden, indem der Lehrende („ t “ für „teacher“) den Lernenden („ l “ für „learner“) belohnt („ G “ für „gratification“) bzw. bestraft („ P “ für „punishment“): $G(t, l)$ und $P(t, l)$.¹¹ Lernen lässt sich dabei als ein Normenschluss verstehen, ausgehend von einer gewissen Bedingungs- bzw. Folgerungsrelation (dargestellt als notwendige bzw. strikte Implikation) zwischen einem bestimmten Verhalten des Lernenden und einer belohnenden oder bestrafenden Lehrendenreaktion. Wir nehmen zudem an, dass der Lernende Belohnungen will: $\mathbf{O}^l(G(t, l))$.¹² Um den deontischen modus ponens anwenden zu können, betrachten wir das (zu verstärkende) Verhalten des Lernenden, $V(l)$, als notwendige Bedingung für die Belohnung:¹³

$$\mathbf{N}(G(t, l) \rightarrow V(l)) \quad (3)$$

Das Erlernete tritt dann im Falle von Belohnungen als seitens des Lernenden gewolltes Verhalten, $\mathbf{O}^l(V(l))$, auf:

$$\mathbf{N}(G(t, l) \rightarrow V(l)) \wedge \mathbf{O}^l(G(t, l)) \rightarrow \mathbf{O}^l(V(l)) \quad (4)$$

Da gilt, dass $p \rightarrow q$ logisch äquivalent mit $\neg q \rightarrow \neg p$ ist, können wir anstelle von (4) auch schreiben:

$$\mathbf{N}(\neg V(l) \rightarrow \neg G(t, l)) \wedge \mathbf{O}^l(G(t, l)) \rightarrow \mathbf{O}^l(V(l)) \quad (5)$$

Umgekehrt gehen wir davon aus, dass der Lernende Bestrafungen nicht will, $\mathbf{O}^l(\neg P(t, l))$, und dass der Lehrende dies zum Lehren des Unterlassens eines Verhaltens verwendet. Formal fassen wir ein (zu unterlassendes) Verhalten $V(l)$ als hinreichende Bedingung für eine Bestrafung auf:

$$\mathbf{N}(V(l) \rightarrow P(t, l)) \quad (6)$$

Mit dem deontischen modus tollens erhalten wir:

$$\mathbf{N}(V(l) \rightarrow P(t, l)) \wedge \mathbf{O}^l(\neg P(t, l)) \rightarrow \mathbf{O}^l(\neg V(l)) \quad (7)$$

Das Erlernen von auszuführendem Verhalten ist in unserer Modellierung also das Ergebnis eines deontischen modus ponens, das Erlernen von zu unterlassendem Verhalten das Resultat eines deontischen modus tollens.¹⁴ Allerdings ist die Ausführung von nicht-gewünschtem Verhalten (logisch) äquivalent mit dem Unterlassen von gewünschtem Verhalten: Das Verhalten

⁹Dass dies für klassisches Konditionieren gilt, wurde bereits 1988 von Robert Rescorla gezeigt [16].

¹⁰Was jedoch eine Belohnung oder Bestrafung ist, ist abhängig davon, ob der Reiz vom jeweiligen Individuum als positiv oder negativ klassifiziert wird. Vgl. für eine formallogische Ausarbeitung [10].

¹¹Dabei nehmen wir an, dass es nicht möglich ist ($= \neg \mathbf{M}(p)$), ein und dasselbe Verhalten zugleich zu belohnen und zu bestrafen: $\neg \mathbf{M}(G(t, l) \wedge P(t, l))$. Zugleich sehen wir Belohnen und Bestrafen nicht in einer kontradiktorischen, sondern einer konträren Relation, was im Vergleich zum traditionellen operanten Konditionieren die Variante eröffnet, sowohl nicht zu belohnen als auch nicht zu bestrafen.

¹²Der hochgestellte kleine Buchstabe hinter dem \mathbf{O} -Operator markiert die Zugehörigkeit, die Perspektivität, des Wollens zum mit diesem Kleinbuchstaben symbolisierten Individuum.

¹³Als notwendig können Sachverhalte angesehen werden, wenn sie den Nachsatz einer Subjunktion bilden. Als hinreichend gelten Sachverhalte, wenn sie den Vordersatz einer Subjunktion bilden. Im Falle von $p \rightarrow q$ gilt somit q als notwendige Bedingung für p und p als hinreichende Bedingung für q . Vgl. hierzu auch [7].

¹⁴Dass Belohnen in einer notwendigen, Bestrafen hingegen in einer hinreichenden Relation zum Verhalten steht, ist die Folge des bis dahin entwickelten Formalismus, konkret des deontischen modus ponens bzw. deontischen modus tollens.

des Lernenden stellt jeweils eine Abweichung vom vom Lehrenden gewünschten bzw. nicht-gewünschten Verhalten (= non-konformes Verhalten) dar. Daher kann (7) auch wie folgt geschrieben werden:

$$\mathbf{N}(\neg V(l) \rightarrow P(t,l)) \wedge \mathbf{O}^l(\neg P(t,l)) \rightarrow \mathbf{O}^l(V(l)) \quad (8)$$

Um dem Lernenden beizubringen, dass er ein bestimmtes Verhalten $V(l)$ zeigen soll, muss der Lehrende also lediglich dafür sorgen, dass wenigstens einer der Vordersätze von (5) oder (8) erfüllt ist. Es genügt also, zu gewährleisten, dass die folgende Formel gilt:

$$\mathbf{N}(\neg V(l) \rightarrow P(t,l)) \vee \mathbf{N}(\neg V(l) \rightarrow \neg G(t,l)) \quad (9)$$

Will der Lehrende, dass der Lernende das Verhalten V zeigt, $\mathbf{O}^l(V(l))$, so muss er im Falle der Unterlassung dieses Verhaltens (also bei non-konformen Verhalten) seitens des Lernenden, $\neg V(l)$, den Lernenden bestrafen und/oder nicht belohnen:¹⁵

$$\mathbf{N}(\mathbf{O}^l(V(l)) \wedge \neg V(l) \rightarrow P(t,l) \vee \neg G(t,l)) \quad (10)$$

Was also zunächst ohne Belang ist, ist eine Reaktion seitens des Lehrenden im Falle konformen Verhaltens, sprich eine Nicht-Bestrafung und Belohnung:

$$\mathbf{N}(\mathbf{O}^l(V(l)) \wedge V(l) \rightarrow \neg P(t,l) \wedge G(t,l)) \quad (11)$$

3. Deduktion von Bedingungsrelationen

Zur Veranschaulichung sei ein Universum der Rede angenommen, in dem der Lehrende zwar das non-konforme Verhalten des Lernenden, $\neg V(l)$, konsequent nicht belohnt, das konforme Verhalten des Lernenden, $V(l)$, aber gerade nicht konsequent belohnt: Ein solcher Fall wäre mit dem folgenden Minimalszenario beschrieben: Der Lernende hat zum Zeitpunkt $k = 0$ das Verhalten V nicht ausgeführt und wurde vom Lehrenden nicht belohnt. Zum Zeitpunkt $k = 1$ hat der Lernende das Verhalten V ausgeführt und wurde vom Lehrenden nicht belohnt. Zum Zeitpunkt $k = 2$ hat der Lernende das Verhalten V nochmals ausgeführt und wurde nun vom Lehrenden belohnt. Insgesamt ergibt sich für das obige Szenario folgende symbolische Übersicht:

$$k = 0 : \neg V(l) \wedge \neg G(t,l); \quad k = 1 : V(l) \wedge \neg G(t,l); \quad k = 2 : V(l) \wedge G(t,l).$$

Da die einzelnen Zeitpunkte eines Universums der Rede im Sinne der ontischen Modallogik¹⁶ als mögliche Welten gedacht werden können, liegt alles, was zu keinem Zeitpunkt vorliegt, unmöglich vor:¹⁷ hier der zu keinem Zeitpunkt (in keiner möglichen Welt) eingetroffene Sachverhalt, dass der Lernende das Verhalten V unterlassen hat und vom Lehrenden belohnt wurde: $\neg \mathbf{M}(\neg V(l) \wedge G(t,l))$. Dies ist jedoch mit der Notwendigkeit, dass nicht gilt, dass der Lernende das Verhalten V unterlassen hat und vom Lehrenden belohnt wurde, logisch äquivalent: $\neg \mathbf{M}(\neg V(l) \wedge G(t,l)) \equiv \mathbf{N}(\neg(\neg V(l) \wedge G(t,l)))$. Notwendig nicht-vorliegenden Konjunktionen sind wiederum logisch äquivalent mit notwendigen Subjunktionen, $\mathbf{N}(\neg(p \wedge q)) \equiv \mathbf{N}(p \rightarrow \neg q)$, hier: $\mathbf{N}(\neg(\neg V(l) \wedge G(t,l))) \equiv \mathbf{N}(\neg V(l) \rightarrow \neg G(t,l))$. Und da $\mathbf{N}(\neg V(l) \rightarrow \neg G(t,l))$ logisch äquivalent mit $\mathbf{N}(G(t,l) \rightarrow V(l))$ ist, hat man (bzw. der Lernende) die für den Normenschluss auf

¹⁵Eine Nicht-Bestrafung und Belohnung von non-konformem Verhalten ist damit ausgeschlossen. Und da wir grundlegend annehmen, dass es nicht möglich ist, ein und dasselbe Verhalten zugleich zu belohnen und zu bestrafen, $\mathbf{M}(G(t,l) \wedge P(t,l))$, ist als Reaktion auf ein non-konformes Verhalten letztlich entweder eine Bestrafung und Nicht-Belohnung oder eine Nicht-Bestrafung und Nicht-Belohnung möglich.

¹⁶Vgl. im Allgemeinen wie auch im Speziellen zur ontischen Modallogik ausführlich [17]. Zur Einführung empfiehlt sich [7]. Aus Perspektive der Informatik ist ein Einstieg mit [18] sinnvoll.

¹⁷Vgl. zur medien- und ingenieurwissenschaftlichen Rechtfertigung dieser Annahme [10].

$\mathbf{O}(V(l))$ notwendige Bedingungsrelation, dass das zu verstärkende Verhalten eine notwendige Bedingung für die Belohnung darstellt, deduziert.¹⁸

Dasselbe gilt nun auch für die Ableitung von $\mathbf{N}(\neg V(l) \rightarrow P(t,l))$ und somit den Fall, dass der Lehrende das non-konforme Verhalten konsequent bestraft, gleichzeitig aber nicht-konsequent konformes Verhaltens nicht bestraft

$$k = 0 : \neg V(l) \wedge P(t,l); \quad k = 1 : V(l) \wedge P(t,l); \quad k = 2 : V(l) \wedge \neg P(t,l).$$

Wiederum ist lediglich entscheidend, dass auf diese Weise $\neg V(l) \wedge \neg P(t,l)$ zu keinem Zeitpunkt vorliegt, also $\neg \mathbf{M}(\neg V(l) \wedge \neg P(t,l))$ gilt, was logisch äquivalent mit $\mathbf{N}(\neg V(l) \rightarrow P(t,l))$ ist. Auch wenn damit im Falle konformen Verhaltens keine konsequente Belohnung und Nicht-Bestrafung notwendig wird, heißt das nicht, dass jegliche Reaktion möglich ist. So sei der Fall angenommen, dass das non-konforme Verhalten des Lernenden (wie gehabt) zum Zeitpunkt $k = 0$ bestraft wird, sein konformes Verhalten aber nicht nur zum Zeitpunkt $k = 1$ sondern auch zum Zeitpunkt $k = 2$ bestraft wird:

$$k = 0 : \neg V(l) \wedge P(t,l); \quad k = 1 : V(l) \wedge P(t,l); \quad k = 2 : V(l) \wedge P(t,l).$$

Wiederum gilt, dass derart $\neg V(l) \wedge \neg P(t,l)$ zu keinem Zeitpunkt vorliegt, also $\neg \mathbf{M}(\neg V(l) \wedge \neg P(t,l))$ als gegeben anzusehen ist, was aufgrund logischer Äquivalenz $\mathbf{N}(\neg V(l) \rightarrow P(t,l))$ entspricht. Allerdings liegt ebenso $V(l) \wedge \neg P(t,l)$ zu keinem Zeitpunkt vor, sodass $\neg \mathbf{M}(V(l) \wedge \neg P(t,l))$ anzunehmen und damit auch $\mathbf{N}(V(l) \rightarrow P(t,l))$ als vorliegend angenommen werden muss. Da wir weiterhin annehmen, dass der Lernende nicht bestraft werden will, $\mathbf{O}^l(\neg P(t,l))$, folgt nach deontischem modus tollens einerseits aufgrund von $\mathbf{N}(\neg V(l) \rightarrow P(t,l))$ das Wollen des konformen Verhaltens, $\mathbf{O}(V(l))$, andererseits aufgrund von $\mathbf{N}(V(l) \rightarrow P(t,l))$ das Wollen des non-konformen Verhaltens, $\mathbf{O}(\neg V(l))$, und damit ein Widerspruch. Vom Lehrenden gewünschtes Verhalten darf also nicht konsequent bestraft werden:

$$\neg \mathbf{N}(V(l) \rightarrow P(t,l)) \tag{12}$$

Entsprechendes gilt auch bezüglich des belohnenden Verhaltens des Lehrenden. Angenommen der Lehrende würde nicht nur (regelentsprechend) das non-konforme Verhalten des Lernenden nicht belohnen, sondern stets auch dessen konformes Verhalten,¹⁹ so würde neben $\mathbf{N}(\neg V(l) \rightarrow \neg G(t,l))$ auch $\mathbf{N}(V(l) \rightarrow \neg G(t,l))$ gelten.²⁰ Das wiederum würde, da wir davon ausgehen, dass der Lernende Belohnungen will, (aufgrund des deontischen modus tollens) sowohl zur Ableitung von $\mathbf{O}(V(l))$ als auch $\mathbf{O}(\neg V(l))$ führen: der Intention des Lernenden nach konformem und non-konformem Verhalten. Der Lehrende darf also konformes Verhalten ebenfalls nicht konsequent nicht belohnen:

$$\neg \mathbf{N}(V(l) \rightarrow \neg G(t,l)) \tag{13}$$

Wir können zusammenfassend feststellen, dass einerseits nur non-konformes Verhalten konsequent via Bestrafung oder Nicht-Belohnung zu sanktionieren ist, konformes Verhalten hingegen nicht konsequent der Belohnung oder Nicht-Bestrafung bedarf.²¹ Andererseits ist der Lehrende hinsichtlich seiner Reaktion auf konformes Verhalten des Lernenden nicht vollkommen frei: Er muss konformes Verhalten entweder konsequent belohnen oder nicht-bestrafen oder aber zumindest konsequent inkonsequent sein, in dem Sinne, dass er konformes Verhalten mal belohnt und nicht-belohnt, mal bestraft und nicht-bestraft.

¹⁸Vgl. zu diesem Vorgehen [10].

¹⁹Bezogen auf das hier stets angeführte Universum der Rede würde gelten: $k = 1 : \neg V(l) \wedge \neg G(t,l)$, $k = 1 : V(l) \wedge \neg G(t,l)$, $k = 3 : V(l) \wedge \neg G(t,l)$.

²⁰ $\mathbf{N}(\neg V(l) \rightarrow \neg G(t,l))$ ist das Resultat dessen, dass zu keinem Zeitpunkt $\neg V(l) \wedge G(t,l)$ und somit $\neg \mathbf{M}(\neg V(l) \wedge G(t,l))$ gilt, was logisch äquivalent mit $\mathbf{N}(\neg V(l) \rightarrow \neg G(t,l))$ ist. $\mathbf{N}(V(l) \rightarrow \neg G(t,l))$ ist hingegen mit $\neg \mathbf{M}(V(l) \wedge G(t,l))$ logisch äquivalent, was daraus folgt, dass zu keinem Zeitpunkt $V(l) \wedge G(t,l)$ gilt.

²¹Wenn auch ein solches nicht schadet, in dem Sinne, dass es nicht nicht möglich/erlaubt wäre.

4. Skinners operantes Konditionieren

Das operante Konditionieren nach Skinner lässt sich mittels eines ABC-Schemas, einem Diagramm der Form

$$\text{Antecedents} \longrightarrow \text{Behavior} \longrightarrow \text{Consequences}, \quad (14)$$

veranschaulichen. In der psychologischen Forschung assoziiert man zu Antecedents, Behavior, und Consequences jeweils Beobachtungsdaten, um dann darauf basierende statistische Korrelationen zu messen. Hier geht es jedoch nicht um statistische Daten, sondern um eine formallogische Analyse von Lernprozessen, die auf operantem Konditionieren aufbauen. Deshalb ordnen wir den Bestandteilen des ABC-Schemas formallogische Formeln zu, die sich von Zeitpunkt zu Zeitpunkt ändern können. Wir assoziieren zu jedem Zeitpunkt k ein ABC-Schema

$$A^k \longrightarrow B^k \longrightarrow C^k \quad (15)$$

und assoziieren zu A^k ein Protokoll der Ereignisse vor dem Zeitpunkt k , zu B^k das Verhalten $b^k(l) \in \{V(l), \neg V(l)\}$ des Lernenden zum Zeitpunkt k , und zu C^k das darauffolgende Verhalten $c^k(t) \in \{G(t,l), \neg G(t,l), P(t,l), \neg P(t,l)\}$ des Lehrenden; formal:

$$\begin{aligned} A^k &: \{b^i(l) \wedge c^i(t) : 0 \leq i < k\}, \\ B^k &: b^k(l), \\ C^k &: c^k(t). \end{aligned}$$

Die Menge der aussagenlogischen Formeln, die zu A^k assoziiert sind, beschreibt diejenigen möglichen Welten, die vor dem Zeitpunkt k vom Lernenden beobachtet wurden. Nach dem Zeitpunkt k wird dann diese Menge um eine neue mögliche Welt $b^k(l) \wedge c^k(t)$ erweitert, weswegen dann die Formeln mit Möglichkeits- oder Notwendigkeitsoperatoren neu ausgewertet werden müssen. Die Anzahl der für einen erfolgreichen Lernprozess erforderlichen Zeitpunkte hängt von der Reaktion des Lehrenden auf konformes Verhalten des Lernenden ab. Mit (12) und (13) sind die Bedingungen beschrieben, die dahingehend gegeben sein müssen:²²

$$\mathbf{M}(V(l) \wedge \neg P(t,l)) \quad (16)$$

$$\mathbf{M}(V(l) \wedge G(t,l)) \quad (17)$$

Es muss also mindestens eine mögliche Welt, ein Zeitpunkt, vorliegen, in dem konformes Verhalten einerseits nicht bestraft und andererseits belohnt wird. Andernfalls kommt es aufgrund der besagten Normenschlüsse zum deontischen und damit logischen Widerspruch $\mathbf{O}(V(l)) \wedge \mathbf{O}(\neg V(l))$.²³ Unabhängig davon erlaubt unsere Modellierung, zum frühestmöglichen Zeitpunkt den Lernprozess zu beenden: dann also, wenn unter Bedingung von (16) und (17) $\mathbf{O}(V(l))$ deduziert wurde.

5. Ereignismodell und operantes Konditionieren

Der sowjetische Literatur- und Kulturwissenschaftler Jurij M. Lotman fasst Äußerungen erst dann als narrativ, also als ereignishaft, auf, wenn „ein Faktum [...] stattgefunden hat, obwohl es nicht hätte stattfinden sollen“ [19]. Auch wenn Lotman lediglich metasprachlich zwischen

²²(12) ist logisch äquivalent mit (16), (13) logisch äquivalent mit (17).

²³Vgl. zur Struktur eines deontischen Widerspruchs im Vergleich zu einem normativen Widerspruch [7].

Soll- und Istsätzen unterscheidet, kann seine Definition von (narrativen) Ereignissen mithilfe des Gebotsoperators der deontischen Modallogik zufriedenstellend formalisiert werden [11]:

$$\mathbf{O}(p) \wedge \neg p \leftrightarrow E \quad (18)$$

Während $\mathbf{O}(p)$ für den Sollzustand und $\neg p$ für den Istzustand steht, steht E für das Vorliegen eines Ereignisses, sodass die Bisubjunktion wie folgt zu lesen ist: „Ein Ereignis liegt genau dann vor, wenn das Gesollte/Gewollte ($= p$ in $\mathbf{O}(p)$) in kontradiktorischer Relation zum Gegebenen ($= \neg p$) steht“. Mithilfe dieser Bisubjunktion kann sodann ausgehend vom jeweiligen Soll- und Istzustand das Vorliegen oder Nicht-Vorliegen von Ereignissen deduziert werden:

$$\mathbf{O}(p) \wedge \neg p \rightarrow E \quad (19)$$

$$\mathbf{O}(p) \wedge p \rightarrow \neg E \quad (20)$$

Ebenso möglich ist der Schluss auf den zugrundeliegenden Sollzustand. Da ein Ereignis definitionsgemäß die Verletzung eines Gesollten oder Gewollten ist und damit eine Abweichung (einen Verstoß) darstellt, wird es als solches in der Narration markiert [20]. Ausgehend vom Vorliegen derartiger Markierungen kann vom Vorliegen eines Ereignisses ausgegangen werden und davon ausgehend dann auf den entsprechenden Sollzustand, also die Gültigkeit eines bestimmten Sollsatzes geschlossen werden:

$$E \wedge \neg p \rightarrow \mathbf{O}(p) \quad (21)$$

Umgekehrt kann ausgehend vom Nicht-Vorhandensein derartiger Markierungen angenommen werden, dass bezüglich des jeweiligen Istzustandes kein Ereignis vorliegt und wiederum mithilfe der Ereignisbisubjunktion die Nicht-Gültigkeit eines bestimmten Sollsatzes deduziert werden:

$$\neg E \wedge \neg p \rightarrow \neg \mathbf{O}(p) \quad (22)$$

Vergleicht man die Schlussmöglichkeiten von Lotmans Ereignismodell mit denen des operanten Konditionierens, so fallen relevante Gemeinsamkeiten auf: Beide basieren auf Sollsätzen bzw. erlauben deren Ableitung. Die Schlussverfahren (19) und (20), also Schlüsse auf den jeweiligen Ereigniszustand ausgehend von einem gegebenen Soll- und Istzustand, können analog zu den Implikationen (10) und (11) gesehen werden, den ‚Regeln‘ für die Reaktion des Lehrenden auf ein konformes oder non-konformes Verhalten des Lernenden: Auch hier wird ausgehend von einem Sollsatz (vom Lehrenden gewünschtes Verhalten des Lernenden) und einem Istsatz (dem konformen oder non-konformen Verhalten des Lernenden) auf Bestrafung und/oder Nicht-Belohnung bzw. Nicht-Bestrafung und Belohnung geschlossen. Damit entspräche das Ereignis auch strukturell dem, was eine Bestrafung bzw. genauer eine Bestrafung und/oder Nicht-Belohnung darstellt. Allerdings würden sich so nicht nur die Schlussverfahren (19) und (20) mit den Implikationen (10) und (11) decken, sondern (auch wenn in eingeschränkterer Form) das Schlussverfahren (21) mit den Implikationen (3) und (5): Während in (21) ausgehend vom Vorliegen eines Ereignisses und eines Istzustandes auf einen Sollzustand geschlossen wird, wird in (4) und (7) ebenfalls auf einen Sollzustand, auf das Sollen oder Wollen eines Verhaltens seitens des Lernenden geschlossen.

Das Wollen in (4) und (7) ist jedoch das Ergebnis eines Schlusses ausgehend von einem anderen Wollen, nämlich jenem belohnt und nicht-bestraft zu werden, und einer zuvor erlernten Bedingungs- bzw. Folgerungsrelation zwischen Belohnung und Bestrafung einerseits und einem bestimmten Verhalten andererseits. Und auch das damit implizierte Vorliegen eines Ereignisses entspricht einer bloßen Markierung eines non-konformen Istzustandes, während Bestrafung/Nicht-Bestrafung und Belohnung/Nicht-Belohnung als Konsequenz einer Non-

Konformität zu werten ist.²⁴ Dies spiegelt sich insbesondere in formaler Hinsicht wider: Das Ereignis als solches weist nicht die Struktur einer prädikatenlogischen Aussage (also als Kombination von Prädikat und Individuen) sondern lediglich einer aussagenlogischen Aussage der Form „Es liegt ein Ereignis vor“ bzw. „Es liegt kein Ereignis vor“. Entsprechend besteht der bereits betrachtete Unterschied zwischen dem Schlussverfahren (21) und den Implikationen (4) und (7) darin, dass im Ereignismodell unmittelbar auf das zu Wollende oder Sollende geschlossen wird: ausgehend lediglich von der Markierung der Non-Konformität eines Zustandes bzw. Verhaltens. Derselbe Schluss bedarf im Falle des operanten Konditionierens hingegen des Wissens bzw. der vorherigen Ableitung der Relation zwischen Verhalten und Bestrafung/Belohnung.

Das operante Konditionieren mag zwar derart umständlicher und den ganzen Lehr- und Lernprozess verkomplizierender erscheinen lassen, allerdings handelt es sich dabei (wie bereits erwähnt) um ein wesentliches Merkmal dieser Form des Lernens: Der Lernende muss nicht über die Intention des Lehrenden Bescheid wissen. Das Lernen erfolgt indirekt über Bestrafung und Belohnung. Darüber hinaus ist das, was beim Schlussverfahren im Ereignismodell schlicht als Markierung der Non-Konformität eines Zustands bzw. Verhaltens bezeichnet wurde, im konkreten Kontext weniger eindeutig als es sprachlich erscheint: Im Normalfall wird mehr als nur ein Istzustand, ein Verhalten, existent sein, sodass trotz Vorliegen eines Ereignisses nicht per se gewusst werden kann, welcher Istzustand, welches Verhalten, das Ereignis nach sich gezogen hat [11].²⁵ Entsprechend kann nur auf eine Menge potenzieller Sollzustände (formal einer Adjunktion von einzelnen Sollzuständen) geschlossen werden, von denen man lediglich weiß, dass mindestens einer davon zutrifft. Der Ausschluss von potenziellen Sollzuständen muss dann durch die Deduktion von negierten Sollzuständen nach dem Schlussverfahren (22) erfolgen, allerdings ist das nur bis zu einem bestimmten Grad möglich [11]. Im Kontext des operanten Konditionierens ist eine solche Uneindeutigkeit zumindest bezüglich der analogen Schlussverfahren (4) und (7) nicht gegeben, sodass es auch nicht überrascht, dass es beim operanten Konditionieren kein Analogon zum besagten Schlussverfahren (22) gibt.

Noch ein letzter Vergleich lohnt: Während das Ereignismodell lediglich das Vorhandensein oder Nicht-Vorhandensein von Ereignissen kennt, E oder $\neg E$, weist das operante Konditionieren auf analoger Ebene mit Bestrafen und Nicht-Belohnen, Nicht-Bestrafen und Belohnen, Nicht-Bestrafen und Nicht-Belohnen drei Varianten auf: $P(t,l) \wedge \neg G(t,l)$, $\neg P(t,l) \wedge G(t,l)$, $\neg P(t,l) \wedge \neg G(t,l)$. Die aufgrund von $P(t,l)/\neg P(t,l)$ und $G(t,l)/\neg G(t,l)$ kombinatorisch mögliche vierte Variante, Bestrafen und Belohnen, $P(t,l) \wedge G(t,l)$, fällt als von vornherein semantisch unvereinbar weg, da Belohnen und Bestrafen stets in einem konträren Verhältnis zueinander stehen. Während nun aber aufgrund der Analogie der Schlussverfahren (11) und (20) (in beiden Fällen wird von einem Wollen und einem dazu konformen Sein geschlossen) auch deren Nachsätze, also $\neg P(t,l) \wedge G(t,l)$ und $\neg E$, analog zu sehen sind, ist aufgrund der Analogie von (10) und (19) (in beiden Fällen wird von einem Wollen und einem dazu non-konformen Sein geschlossen) und damit auch deren Nachsätze, E sowohl zu $P(t,l) \wedge \neg G(t,l)$ als auch $\neg P(t,l) \wedge \neg G(t,l)$ analog zu sehen. Das operante Konditionieren ist an dieser Stelle also differenzierter und erlaubt so etwas wie eine starke Bestrafung in Form einer Bestrafung und Nicht-Belohnung oder eine Art schwache Bestrafung in Form einer bloßen Nicht-Belohnung.

²⁴Es sei denn, man möchte die Markierung eines non-konformen Zustandes oder Verhaltens bereits als Sanktionierung (Bestrafung und Nicht-Belohnung) dieses non-konformen Zustandes bzw. Verhaltens ansehen.

²⁵Dieser Sachverhalt spiegelt sich nicht zuletzt darin wider, dass im entsprechenden Schlussverfahren (21) von einem unspezifizierten Ereignis E und einem Istzustand $\neg p$ auf $\mathbf{O}(p)$ geschlossen wird und gerade nicht von einem spezifizierten Ereignis in Form eines Ereignis-Operators \mathbf{E} mit einem $\neg p$ im Wirkungsbereich: $\mathbf{E}(\neg p)$.

Rückblick und Ausblick

In dieser Arbeit haben wir die klassische Verbindung des in der Modellierung von technischen Systemen erfolgreich eingesetzten ABC-Schemas [21, 9] mit dem operanten Konditionieren mittels eines durch die formale Subtextanalyse [11] motivierten formallogischen Kalküls beschrieben. Das ermöglicht eine Automatisierung bestimmter Lernprozesse und ist daher ein wichtiger Schritt zur Konstruktion lernfähiger kognitiver Agenten.

Literatur

- [1] SKINNER, B. F.: *The concept of the reflex in the description of behavior. Journal of General Psychology*, 5, S. 427–458, 1931.
- [2] SKINNER, B. F.: *The Behavior of Organisms: An Experimental Analysis*. B.F. Skinner Foundation, Cambridge, Massachusetts, 1938.
- [3] CHATER, N. und M. OAKSFORD: *The probability heuristics model of syllogistic reasoning. Cognitive Psychology*, 38, S. 191–258, 1999.
- [4] CUMMINS, D.: *Naïve theories and causal deduction. Memory & Cognition*, 23, S. 646–658, 1995.
- [5] HAGMAYER, Y.: *Logik in der psychologie. warum menschen nicht gemäß den gesetzen der logik schlussfolgern*. In P. KLIMCZAK und T. ZOGLAUER (Hrsg.), *Logik in den Wissenschaften*. Mentis, Münster, 2017. Im Erscheinen.
- [6] ZOGLAUER, T.: *Normenkonflikte. Zur Logik und Rationalität ethischen Argumentierens*. Frommann-Holzboog, Stuttgart, 1998.
- [7] ZOGLAUER, T.: *Einführung in die formale Logik für Philosophen*. Vandenhoeck & Ruprecht, Göttingen, 2008.
- [8] MORSCHER, E.: *Normenlogik. Grundlagen – Systeme – Anwendungen*. Mentis, Paderborn, 2012.
- [9] KLIMCZAK, P., M. WOLFF, J. LINDEMANN, C. PETERSEN, R. RÖMER, und T. ZOGLAUER: *Die kognitive Heizung*. In *Elektronische Sprachverarbeitung 2014, Tagungsband der 25. Konferenz, Dresden, 26.–28. März*, S. 89–96. 2014.
- [10] KLIMCZAK, P., R. RÖMER, und M. WOLFF: *Ein beitrag zur formalisierung finales systeme*. In *Elektronische Sprachverarbeitung 2017, Tagungsband der 27. Konferenz, Saarbrücken, 15.–17. März*. 2017. In diesem Band.
- [11] KLIMCZAK, P.: *Formale Subtextanalyse. Kalkülisierung von Narration und Interpretation*. Mentis, Münster, 2016.
- [12] BAYER, K.: *Argument und Argumentation. Logische Grundlagen der Argumentationsanalyse*. Vandenhoeck & Ruprecht, Göttingen, 2007.
- [13] NAESS, A.: *Kommunikation und Argumentation. Eine Einführung in die angewandte Semantik*. Scriptor, Kronberg, 1975.
- [14] ANDERSON, J. R.: *Learning and memory: An integrated approach*. Wiley, New York, 2000.

- [15] HOUWER, J. D.: *The propositional approach to associative learning as an alternative for association formation models. Learning & Behavior*, 37, S. 1–20, 2009.
- [16] RESCORLA, R. A.: *Pavlovian conditioning. it's not what you think it is. American Psychologist*, 43, S. 151–160, 1988.
- [17] HUGHES, G. E. und M. J. CRESSWELL: *A New Introduction to Modal Logic*. Routledge, London, 1996.
- [18] SCHENKE, M.: *Logikkalküle in der Informatik: Wie wird Logik vom Rechner genutzt?* Springer, Wiesbaden, 2013.
- [19] LOTMAN, J. M.: *Die Struktur literarischer Texte*. Fink, München, 1993.
- [20] KRAH, H.: *Einführung in die Literaturwissenschaft/Textanalyse*. Ludwig, Kiel, 2006.
- [21] WIRSCHING, G., M. HUBER, C. KÖLBL, R. LORENZ, und R. RÖMER: *Semantic dialogue modeling*. In A. ESPOSITO, A. M. ESPOSITO, A. VINCIARELLI, R. HOFFMANN, und V. C. MÜLLER (Hrsg.), *Behavioral Cognitive Systems*, Bd. 7403 d. Reihe LNCS, S. 104–113. 2012.