

# EIN BEITRAG ZUR FORMALISIERUNG FINALER SYSTEME

Peter Klimczak<sup>1</sup>, Ronald Römer<sup>1</sup>, Matthias Wolff<sup>1</sup>

<sup>1</sup>Brandenburgische Technische Universität Cottbus-Senftenberg  
peter.klimczak@b-tu.de

**Kurzfassung:** Norbert Bischof subsumiert kognitive Agenten unter die finalen Systeme. Als solche müssen *technische* kognitive Agenten situationsabhängig adäquate Aktionen ausführen und diese durch Beobachtung ihrer Umwelt erlernen. Sie müssen also wissen, was sie jeweils tun *sollen*. Es wird im Folgenden gezeigt, dass solche situationsspezifische Präskriptionen unter Zuhilfenahme erweiterter Schlusserfahren auch aus wenigen elementaren, unspezifischen Präskriptionen deduziert werden können. Neben allgemeinen Ausführungen zur formallogischen Repräsentation von Deskriptionen, Präskriptionen und Intentionen, wird ein Konzept zur Deduktion von Bedingungs- und Folgerungsrelationen aus der Korrelation von Agentenverhalten und Umweltreaktion gezeigt und die Möglichkeit erläutert, mithilfe von Bedingungs- und Folgerungsrelationen auf Präskriptionen zu schließen. Die Formalisierung von Lernen als operantes Konditionieren erlaubt zudem die Modellierung und Berücksichtigung einer dynamischen, strafenden und belohnenden Umwelt, die selbst als kognitiver Agent handelt. Zur Veranschaulichung wird das auf den Mathematiker und Elektrotechniker Claude E. Shannon zurückgehende Maus-Labyrinth-Beispiel adaptiert.

## 1. Einführung

Kognitive Agenten können aus Sicht der theoretischen Psychologie in die Klasse der finalen Systeme fallen [1]. Sie müssen situationsabhängig adäquate Aktionen ausführen und diese erlernen können. Dabei gilt für kognitive Agenten (=m) grundlegend die Verhaltensvorschrift (=O(p)), dass, falls die Ausführung eines bestimmten Verhaltens (V(m)) zum Zeitpunkt k gesollt ist, sie zum Zeitpunkt k+1 ausgeführt wird [2]:

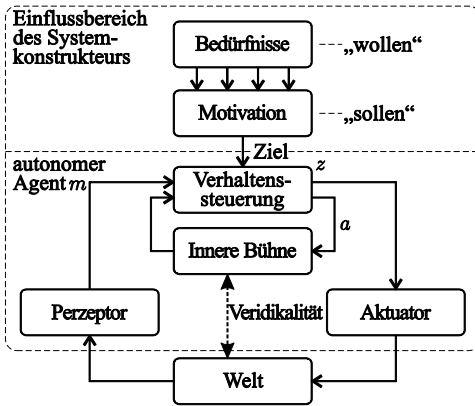
$$(1) \quad \mathbf{O}(V(m))^k \rightarrow V(m)^{k+1}$$

Es muss also ermittelt werden, welches Verhalten in der aktuellen Situation gesollt ist. In technischen kognitiven Systemen geschieht dies in der Regel mit Hilfe von Entscheidungsprozessen, deren Parameter aus Interaktionen mit der Umwelt (=Aktion ausführen und daraufhin Beobachtung sowie Belohnung erhalten) – also aus Deskriptionen (=Daten) – gelernt werden [3] [4]. Allerdings folgen aus Deskriptionen nie Präskriptionen.<sup>1</sup> Daher *müssen* für jedes kognitive System bestimmte Präskriptionen *gesetzt* werden. Wir werden zeigen, dass situationsspezifische Präskriptionen aus wenigen elementaren, unspezifischen Präskriptionen, welche in aktuellen Systemen implizit durch die Konstruktion vorgegeben sind,<sup>2</sup> formallogisch deduziert werden können. Dabei können elementare Präskriptionen als ‚innere Bedürfnisse‘ (oder schwächer: Intentionen) des kognitiven Agenten interpretiert werden. Wir folgen mit unserem Ansatz im Prinzip Dörner [8], nach dem sich der Autonomiegrad eines technischen kognitiven Systems steigern lässt, wenn zwei weitere, psychologisch motivierte, Systemebenen berücksichtigt werden: die Ebene der Motivation und die Ebene der Bedürfnisse (Bild 1).

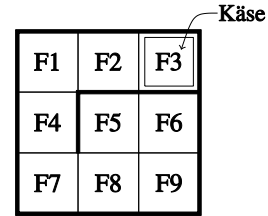
---

<sup>1</sup>Bekannt als „Humes Gesetz“. Vgl. hierzu neben den Original-Ausführungen von Hume [5] auch [6] und [7].

<sup>2</sup>‚Gesollt‘ ist die Aktion, die langfristig die größte Belohnung verspricht. Die Auswahl erfolgt dabei durch den Entscheidungsprozess.



**Bild 1** - Zur Unterscheidung der Einflussbereiche des autonomen Agenten  $m$  und des Systemkonstruktors. Welt ist in unserem Kontext die reale Welt und die innere Bühne die erzählte Welt.



**Bild 2** - Beispielwelt: Einfaches Labyrinth in dem sich der kognitive Agent (Maus) bewegt. Fette Linien kennzeichnen nicht passierbare Wände, Käse befindet sich im Feld F3.

Zur Veranschaulichung adaptieren wir das auf den US-Amerikaner C. E. Shannon zurückgehende Maus-Labyrinth-Beispiel.<sup>3</sup> Der technische kognitive Agent  $m$ , in unserem Beispiel eine Maus, kann Aktionen  $a$  aus einem Aktionsrepertoire ausführen. Im Falle der Maus ist es die Bewegung in eine der vier Haupthimmelsrichtungen,  $a \in \{N, O, S, W\}$ , was wir wie folgt als Prädikat notieren:

$$(2) \quad a(m) \in \{N(m), O(m), S(m), W(m)\}$$

Zu jedem Zeitpunkt  $k$  befindet sich der Agent in einem bestimmten Zustand  $z$ . In unserem Beispiel stellt das in Bild 2 dargestellte Labyrinth die Welt mit den Zuständen  $z \in \{F1, \dots, F9\}$  dar. Die einzelnen Zustände bezeichnen hier das Feld, in dem sich die Maus momentan befindet. Wir notieren den aktuellen Zustand wiederum als Prädikat:

$$(3) \quad z(m) \in \{F1(m), \dots, F9(m)\}$$

Abhängig vom Feld, in dem sich die Maus befindet, kommt sie in den Besitz von Käse oder nicht. Im in Bild 2 dargestellten Labyrinth bekommt die Maus genau dann Käse, wenn sie sich auf dem Feld 3 befindet:

$$(4) \quad F3(m) \leftrightarrow K(m)$$

Wir definieren Verhalten allgemein dahingehend, dass der kognitive Agent eine bestimmte Aktion bei Vorliegen eines bestimmten Weltzustands ausführt, hier das Verhalten, dass die Maus sich in einem der neun Felder befindet und sich in eine der vier Himmelsrichtungen bewegt:

$$(5) \quad V_{z,a}(m) : \leftrightarrow z(m) \wedge a(m)$$

Um in den Besitz von Käse zu gelangen, bedarf es demnach ausgehend vom Feld 2 der Bewegung nach Osten:

$$(6) \quad V_{F2,O}(m) \leftrightarrow F2(m) \wedge O(m)$$

Im weiteren Verlauf symbolisieren wir im Allgemeinen das Verhalten kognitiver Agenten kurz als  $V(m)$ :

$$(7) \quad V(m) : \leftrightarrow V_{z,a}(m)$$

<sup>3</sup>Vgl. zum Beispiel auch [9] in diesem Band.

Das zielführende Verhalten der Maus wird hingegen weiterhin als  $V_{F2,O}(m)$  symbolisiert. Jegliches andere (und damit notwendigerweise nicht zielführende) Verhalten der Maus, kann dementsprechend unter die Negation des zielführenden Verhaltens subsumiert werden:

$$(8) \quad \neg V_{F2,O}(m) \leftrightarrow \neg(F2(m) \wedge O(m))$$

Aufbauend auf dem Beschriebenen sei das folgende Szenario angenommen: Die Maus, auf dem Feld 2 stehend, bewegt sich zum Zeitpunkt  $k=0$  Richtung Westen (zeigt also nicht das zielführende Verhalten) und kommt (infolgedessen) nicht in den Besitz von Käse. Zum Zeitpunkt  $k=1$  bewegt sich die Maus Richtung Osten, da sie aber auf Feld 1 stand, hat sie auch in diesem Fall nicht das zielführende Verhalten gezeigt. Entsprechend ist sie auch weiterhin nicht im Besitz von Käse. Zum Zeitpunkt  $k=2$  bewegt sich die Maus nochmals in Richtung Osten, nunmehr aber vom Feld 2 aus, sodass sie auf das Feld 3 gelangt und in den Besitz von Käse kommt. Insgesamt ergibt sich für die obige Beobachtung folgende symbolische Übersicht:  $k=0$ :  $\neg V_{F2,O}(m) \wedge \neg K(m)$ ;  $k=1$ :  $\neg V_{F2,O}(m) \wedge \neg K(m)$ ;  $k=2$ :  $V_{F2,O}(m) \wedge K(m)$ .

Um ausgehend davon formal-logisch korrekte Schlüsse auf die Ontologie des Universums der Rede schließen zu können, empfiehlt sich eine modallogische Interpretation der medienwissenschaftlichen Vorstellung von erzählten bzw. dargestellten Welten.

## 2. Erzählte Welten und Kausalitäten

In der Medienwissenschaft geht man davon aus, dass Erzählungen nicht die Realität, also die Welt als solche, abbilden, sondern eine eigene Realität, eine ‚eigene‘ Welt, modellieren [10]: die dargestellte oder erzählte Welt. Obwohl die erzählte Welt aufgrund ihrer im Verhältnis zur realen Welt notwendigen Begrenztheit maximal nur einen Ausschnitt der realen Welt abbilden kann, ist die erzählte Welt nicht nur als ein derartiger Ausschnitt der realen Welt aufzufassen, sondern als Vorstellung, als Modell dieser [10].<sup>4</sup> Damit geht einher, dass alles, was in einer erzählten Welt gegeben oder nicht gegeben ist, als ebenso gegeben oder nicht gegeben anzunehmen ist – unabhängig davon, ob es in der realen Welt gegeben oder nicht gegeben ist.

Diese Vorstellung lässt sich adäquat in eine formallogische Modellierung überführen, wenn angenommen wird, dass mit jedem neuen Zeitpunkt und damit jeder Prädikatsänderung der erzählten Welt eine neue mögliche Welt, im Sinne der ontischen Modallogik [11], angenommen wird. Entsprechend dem gewählten System gilt eine Reihe von Gesetzen, die das Verhältnis des Seins ( $=p$ ), des Möglichs ( $=M(p)$ ) und des Notwendigseins ( $=N(p)$ ) in den einzelnen möglichen Welten des Universums der Rede regeln: So liegt alles, was zu keinem Zeitpunkt, in keiner möglichen Welt, vorliegt, bspw.  $p$ , unmöglich,  $\neg M(p)$ , bzw. aufgrund logischer Äquivalenz notwendigerweise nicht vor:  $N(\neg p)$ . Fruchtbar wird diese Erkenntnis jedoch erst dann, wenn nicht nur einzelne unmögliche Aussagen (=ein zu keinem Zeitpunkt vorkommendes Prädikat eines Individuums), sondern unmögliche Aussagenkonjunktionen (=zu keinem Zeitpunkt vorkommende Merkmalskombinationen eines Individuums) betrachtet werden: Aufbauend auf einer rein deskriptiven Erfassung von Individuen (hier kognitive Agenten, in den Medienwissenschaften wiederum Figuren) und ihrer Prädikate (hier ihr Verhalten, in den Medienwissenschaften die Figurenmerkmale) in den einzelnen möglichen Welten werden so nämlich ‚Kausalitäten‘ deduzierbar. Das Besondere daran ist, dass eine solche Deduktion keiner Berücksichtigung der Zeit bedarf: Zwar werden aufgrund von Prädikatswechsel neue Zeitpunkte und damit neue mögliche Welten erfasst, deren Reihenfolge bzw. Ordnung jedoch, bedingt durch das gewählte Kalkül (S5), *nicht* berücksichtigt wird [11] [12].

---

<sup>4</sup>Wie in [9] bezeichnen wir dieses innere Abbild der realen Welt in kognitiven Agenten als Innere Bühne (siehe Bild 1). Der Begriff „Veridikalität“ beschreibt das Maß, in dem das innere Abbild der Welt realitätsgerecht ist [1].

Der Anschaulichkeit wegen sei wieder auf das konkrete, oben beschriebene Maus-Labyrinth-Szenario zurückgegriffen: Der zu keinem Zeitpunkt (in keiner möglichen Welt) eingetretene Sachverhalt, dass die Maus in den Besitz von Käse gekommen ist, ohne (zuvor) das erforderliche Verhalten zu zeigen,  $\neg \mathbf{M}(K(m) \wedge \neg V_{F2,O}(m))$ , ist mit der Notwendigkeit, dass nicht gilt, dass die Maus das entsprechende Verhalten nicht gezeigt hat und im Besitz von Käse ist,  $\mathbf{N}(\neg(\neg V_{F2,O}(m) \wedge K(m)))$ , logisch äquivalent. Notwendig nicht vorliegenden Konjunktionen sind wiederum logisch äquivalent mit notwendigen Implikationen,  $\mathbf{N}(\neg(p \wedge q)) \leftrightarrow \mathbf{N}(p \rightarrow \neg q)$ , hier:  $\mathbf{N}(\neg(\neg V_{F2,O}(m) \wedge K(m))) \leftrightarrow \mathbf{N}(\neg V_{F2,O}(m) \rightarrow \neg K(m))$ . Notwendige (oder *strikte*) Implikationen repräsentieren jedoch nichts weniger als Bedingungs- und Folgerungsrelationen,<sup>5</sup> sodass derart eine exakte und jenseits sprachlicher Ex- und Implikationen und damit verbundener Probleme [13] liegende Ableitung von ‚Kausalitäten‘ möglich wird [14]. Entsprechend ist  $\mathbf{N}(\neg V_{F2,O}(m) \rightarrow \neg K(m))$  zu lesen als: „Wenn die Maus das zielführende Verhalten nicht gezeigt hat, so kommt sie nicht in den Besitz von Käse“. Entscheidend ist, dass diese Relation gewonnen wurde, ohne zuvor zu wissen, dass  $V_{F2,O}(m)$  die notwendige Bedingung für  $K(m)$  ist.<sup>6</sup>

### 3. Intentionen, Präskriptionen, Deskriptionen

Zwar wird von der Philosophie 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 [15]; doch gerade deshalb ist es im hier behandelten Kontext sinnvoll, Intentionen innerhalb der deontischen Modallogik, und damit innerhalb eines Kalküls, das eigentlich nur Präskriptionen zu symbolisieren erlaubt, zu modellieren.<sup>7</sup> Wenn nämlich Widersprüchliches, ohne dass ein formallogischer Widerspruch abgeleitet werden kann, zugleich gewollt werden kann, so sind die deduzierten Wollenssätze ohne Wert: Ihnen kann unmöglich entsprochen werden. Zwar kann Widersprüchliches sehr wohl gewollt, aber keineswegs erfüllt werden. So kann die Maus sowohl das Bedürfnis verspüren in den Besitz von Käse zu gelangen als auch zu kopulieren; wenn das eine aber nicht mit dem anderen realisierbar ist, so bleibt es auch unabhängig vom Wollen beider nicht realisierbar. Der Versuch beides zu realisieren würde ganz im Gegenteil dazu führen, dass weder das eine noch das andere realisiert wird. Da im Kontext von kognitiven Agenten auch Intentionen die Funktion von Verhaltensvorschriften haben, ist es nur folgerichtig, Intentionen wie Präskriptionen zu behandeln und mittels Sollsätzen abzubilden.

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  $\neg p$ , das Gesolltsein von  $\neg p$  folgt,<sup>8</sup> sind keine Schlüsse aus Deskriptionen (Tatsachen) auf Präskriptionen und (in unserem Fall auch) Intentionen möglich und umgekehrt. Aus Deskriptionen allein (und somit dem, was als faktisch bezeichnet werden kann) ergeben sich weder Präskriptionen noch Intentionen.

<sup>5</sup>Zu notwendigen und hinreichenden Bedingungen und deren Darstellung mittels Subjunktionen vgl. [12].

<sup>6</sup>Auch wenn in der obigen metasprachlichen Beschreibung der einzelnen Zeitpunkte in Klammern kausale Relationen angegeben wurden, so geschah das nur des besseren Verständnisses wegen. Das ‚eingeklammerte‘ Wissen ist nicht notwendig, es findet sich auch in der formallogischen Repräsentation der drei einzelnen Zeitpunkte nicht wieder.

<sup>7</sup>Vgl. zur deontischen Logik ausführlich [16]. Zum Einstieg bietet sich hingegen [12] an, da trotz der Kürze nicht nur Probleme der deontischen Logik benannt, sondern auch Vorschläge für deren Lösung gemacht werden. Dass diese Probleme im Kontext medienwissenschaftlicher und kommunikationstechnischer Anwendungen keine Rolle spielen, zeigt wiederum ausführlich [14].

<sup>8</sup>Ganz im Gegensatz zur ontischen Modallogik, in der gilt, dass aus dem Notwendigsein von  $p$ , das Sein von  $p$  folgt:  $\mathbf{N}(p) \rightarrow p$ .

## 4. Deduktion von Präskriptionen/Intentionen

Aus genau diesem Grund sind Folgerungsrelationen und deren Ableitung (wie in Abschnitt 3 beschrieben) so wertvoll, weil sie die Deduktion von Sollsätzen erlauben. Dies allerdings nur unter Hinzunahme bereits bestehender Präskriptionen und Intentionen. Entgegen linguistischen und argumentationstheoretischen Annahmen [17][18] kann bei einem solchen Normenschlussverfahren nicht vom Sollen der Folge auf das Sollen der Ursache, also vom Sollen des Nachsatzes der Subjunktion auf das Sollen des Vordersatzes geschlossen werden. Das würde unweigerlich zu unmittelbaren und mittelbaren Widersprüchen und damit Fehlschlüssen schon in den einfachsten Kontexten führen.<sup>9</sup> Formallogisch möglich, also widerspruchsfrei, ist die Ableitung von mittelbaren Sollsätzen nur in umgekehrter Richtung, 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“) [14]:

$$(9) \quad \mathbf{O}(p) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(q)$$

$$(10) \quad \mathbf{O}(\neg q) \wedge \mathbf{N}(p \rightarrow q) \rightarrow \mathbf{O}(\neg p)$$

Bezogen auf das obige Maus-Labyrinth-Szenario, in dem bekanntlich gilt, dass die Maus nicht in den Besitz von Käse kommt, wenn sie nicht das entsprechende Verhalten zeigt,  $\mathbf{N}(\neg V(m) \rightarrow \neg K(m))$ , kann (nur) im Falle dessen, dass das In-Besitz-Kommen von Käse gesollt oder gewollt ist,  $\mathbf{O}(K(m))$ , via deontischem modus tollens auf das Sollen oder Wollen des zielführenden Verhaltens  $V$  geschlossen werden  $\mathbf{O}(V(m))$ :

$$(11) \quad \mathbf{O}(K(m)) \wedge \mathbf{N}(\neg V(m) \rightarrow \neg K(m)) \rightarrow \mathbf{O}(V(m))$$

Neben diesen (letztlich) beiden einzigen Möglichkeiten neue, also nicht bereits gesetzte oder gegebene, Präskriptionen oder Intentionen abzuleiten, gibt es noch zwei (lediglich funktionale)<sup>10</sup> 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,  $\mathbf{N}(p) \rightarrow \mathbf{O}(p)$ . Beim Ersteren handelt es sich um ein traditionsreiches Rechtsprinzip [20],<sup>11</sup> das auch im Kontext von technischen kognitiven Systemen sinnvoll ist, da Wollensaussagen die Funktion von Verhaltensvorschriften haben: Wenn etwas nicht möglich ist, ist es auch nicht realisierbar, sodass dessen Wollen bzw. die Anweisung es zu realisieren keinen Sinn hat. Allerdings handelt es sich bei  $\neg \mathbf{O}(p)$  um keine Präskription oder Intention im engeren Sinne, sondern lediglich um die Negation einer Präskription bzw. Intention. Dies spiegelt sich auch in der logischen Äquivalenz von  $\neg \mathbf{O}(p)$  mit  $\mathbf{P}(\neg p)$  wieder, der Permission (und damit weder Präskription noch Intention) der Negation dessen, was nicht gewollt oder gesollt ist. Hingegen wird im Falle von  $\mathbf{N}(p) \rightarrow \mathbf{O}(p)$  mit  $\mathbf{O}(p)$  tatsächlich auf eine Präskription/Intention geschlossen – und das ausgehend von einer Deskription, also etwas lediglich ‚Faktischem‘.<sup>12</sup> Obwohl damit das Sein, wenn auch in Form eines

<sup>9</sup>Vgl. hierzu kurz [19] in diesem Band oder ausführlich [14].

<sup>10</sup>Funktional sind sie deshalb, weil beide in komplexeren Kontexten für widerspruchsfreie Schlussverfahren nach deontischem modus ponens und deontischem modus tollens unverzichtbar sind [14]

<sup>11</sup>Entsprechend ist dieses Rechtsprinzip auch als „Impossibilia nulla est obligatio“ (Zur Unmöglichkeit besteht keine Verpflichtung) oder „Ultra posse nemo obligatur“ (Über sein Können hinaus ist niemand verpflichtet) bekannt.

<sup>12</sup>Zwar handelt es sich hierbei um kein traditionsreiches Rechtsprinzip, dennoch stellt diese Relation zwischen Notwendigem und Gebotenem einen in der Rechtsphilosophie diskutierten Sachverhalt dar [20]. Wobei nicht nur die Auffassung, dass das Notwendige, gerade weil es notwendig ist, auch geboten sein muss, vertreten wird, sondern auch die komplett gegensätzliche Auffassung, dass etwas, was notwendig ist, nicht geboten sein kann [20]. Letzteres wäre symbolisch als  $\mathbf{N}(p) \rightarrow \neg \mathbf{O}(p)$  abzubilden, woran wiederum zu sehen ist, dass dies in

Notwendigsein, ein Sollen nach sich zieht, handelt es sich im Kontext kognitiver Agenten, in welchem sowohl Präskriptionen als auch Intentionen die Funktion von Verhaltensvorschriften haben, um eine vollkommen problemlose Relation: Etwas, das notwendig ist, ist als solches schon realisiert, sodass die Vorschrift zu dessen Realisierung als erfüllt gelten muss.

## 5. Operantes Konditionieren und kognitives Lernen

Aus psychologischer Sicht stellt Lernen einen Prozess dar, der als Ergebnis von Erfahrungen relativ langfristige Änderungen im Verhaltenspotential zur Folge hat [21]. Mittels der soeben beschriebenen Schlussverfahren, lässt sich ein Modell entwerfen, dass diese Definition erfüllt und damit Lernen für finale Systeme anwendbar macht,<sup>13</sup> zumindest dann, wenn Lernen als operantes Konditionieren verstanden wird. Dass ein operantes Konditionieren dabei Kognition nicht ausschließt, zeigen nicht zuletzt kognitive Ansätze in der psychologischen Forschung [24].<sup>14</sup> Entscheidend ist beim operanten Konditionieren, dass die Änderung des Verhaltenspotenzials operant oder instrumentell erfolgt: ein bestimmtes Verhalten des Agenten wird seitens der Umwelt reaktiv belohnt oder bestraft mit der Folge, dass die Ausführung bzw. das Unterlassen dieses Verhaltens verstärkt, sprich erlernt, wird. Bestrafung und Belohnung bzw. bestrafender und belohnender Umweltreiz können als zweistellige Prädikate formalisiert werden, indem die Umwelt  $u$  den (kognitiven) Agenten  $m$  belohnt („G“ für „gratification“) bzw. bestraft („P“ für „punishment“):

$$(12) \quad G(u, m)$$

$$(13) \quad P(u, m)$$

Lernen lässt sich demnach als ein (bewusster oder unbewusster) Normenschluss verstehen, ausgehend zum einen von einer gewussten Folgerungsrelation (dargestellt als strikte Implikation) zwischen einem bestimmten Verhalten des Agenten und einem belohnenden oder bestrafenden Umweltreiz und zum anderen dem Wunsch des Agenten nach Belohnung,  $\mathbf{O}^m(G(u, m))$ ,<sup>15</sup> und Nicht-Bestrafung,  $\mathbf{O}^m(\neg P(u, m))$ :

$$(14) \quad N(G(u, m) \rightarrow V(m)) \wedge \mathbf{O}^m(G(u, m)) \rightarrow \mathbf{O}^m(V(m))$$

$$(15) \quad N(V(m) \rightarrow P(u, m)) \wedge \mathbf{O}^m(\neg P(u, m)) \rightarrow \mathbf{O}^m(\neg V(m))$$

Entsprechend dem deontischen modus ponens in (14) und dem deontischen modus tollens in (15) muss davon ausgegangen werden, dass die zu verstärkende Ausführung des Verhaltens eine notwendige Bedingung für die Belohnung darstellt (=16) und das zu verstärkende Unterlassen des Verhaltens eine hinreichende Bedingung für die Belohnung darstellt (=17):

$$(16) \quad N(G(u, m) \rightarrow V(m))$$

---

unserem Kontext kaum von Interesse sein kann, da lediglich auf die Negation einer Präskription oder Intention geschlossen wird und damit weder auf eine Präskription noch Intention.

<sup>13</sup>Wie die empirische Forschung der Psychologie eindringlich gezeigt hat, entspricht das Schließen nach (Standard-)Kalkülen der formalen Logik nicht dem ‚natürlichen‘ Schließen von Menschen. So sehr einzelne alternative Ansätze und Modelle überzeugen mögen [22], so wenig sind sie für eine Verwendung im Kontext technischer kognitiver Agenten geeignet, handelt es sich doch zumeist nicht um in Form von Kalkülen dargelegte Modelle (ausgenommen vielleicht jene Ansätze, die sich bspw. der Default-Logik bedienen, [23]). Allerdings hat der hier vorgestellte Ansatz den großen Vorteil, dass sich die verwendete Prädikatenlogik recht leicht um ontische und deontische Modaloperatoren erweitern lässt (und damit ausdrucksstärker wird).

<sup>14</sup>Für das klassische Konditionieren wurde bereits 1988 im Ansatz gezeigt, dass es sich auf logische Schlussfolgerungen aufbauen lässt [25].

<sup>15</sup>Der hochgestellte kleine Buchstabe hinter dem Gebotsoperator markiert die Zugehörigkeit, die Perspektivität, des Wollens zum mit diesem Kleinbuchstaben symbolisierten Individuum (in diesem Kontext: die Umwelt oder der Agent).

$$(17) \quad N(V(m) \rightarrow P(u, m))$$

Hingegen tritt das Erlernete im Falle von Belohnungen als gewolltes Verhalten (=18) und im Falle von Bestrafungen als nicht gewolltes Verhalten (=19) auf:

$$(18) \quad \mathbf{O}^m(V(m))$$

$$(19) \quad \mathbf{O}^m(\neg V(m))$$

Aufgrund der Schlussverfahren (14) und (15) gilt zudem, dass die Ausführung von der Umwelt nicht gewünschten Verhaltens von der Umwelt zu bestrafen ist (=20), während das Unterlassen von der Umwelt gewünschten Verhaltens von der Umwelt nicht zu belohnen ist (=21):

$$(20) \quad \mathbf{O}^u(\neg V(m)) \wedge V(m) \rightarrow P(u, m)$$

$$(21) \quad \mathbf{O}^u(V(m)) \wedge \neg V(m) \rightarrow \neg G(u, m)$$

Die Ausführung von nicht gewünschtem Verhalten ist jedoch (logisch) äquivalent mit dem Unterlassen von gewünschtem Verhalten: Das Verhalten des Agenten stellt jeweils eine Abweichung vom von der Umwelt gewünschten bzw. nicht gewünschten Verhalten dar (= non-konformes Verhalten). Entsprechend sind aufgrund der lediglich zwei gültigen Normenschlussregeln (dem deontischen modus ponens und modus tollens) nur Bestrafungen oder Nicht-Belohnungen nötig und somit nur im Falle non-konformen Verhaltens eine Reaktion seitens der Umwelt nötig.<sup>16</sup>

$$(22) \quad \mathbf{O}^u(V(m)) \wedge \neg V(m) \rightarrow P(u, m) \vee \neg G(u, m)$$

Ohne Belang ist also eine Reaktion seitens der Umwelt im Falle konformen Verhaltens, sprich eine Nicht-Bestrafung und Belohnung:

$$(23) \quad \mathbf{O}^u(V(m)) \wedge V(m) \rightarrow \neg P(u, m) \wedge G(u, m)$$

Zur Veranschaulichung sei nun ein abgewandeltes, dynamisches Maus-Käse-Szenario angenommen. Für alle Zeitpunkte gilt dabei, dass analog (22) die Ausführung des Verhaltens V gewünscht wird (=24), die Unterlassung des Verhaltens V nicht belohnt wird (=25) und die Nicht-Belohnung das Nicht-in-Besitz-Kommen von Käse darstellt (=26):

$$(24) \quad \mathbf{O}^u(V(m))$$

$$(25) \quad \mathbf{O}^u(V(m)) \wedge \neg V(m) \rightarrow \neg G(u, m)$$

$$(26) \quad \neg G(u, m) \leftrightarrow \neg K(m)$$

Ferner gehen wir davon aus, dass unsere (kognitive) Maus weder (24) noch (25) weiß, dafür aber im Stande ist, ihr Verhalten und das der Umwelt wahrzunehmen und die zuvor beschriebenen Deduktionen vorzunehmen. Im nun betrachteten Szenario hat die Maus zum Zeitpunkt  $k=0$  das seitens der Umwelt gewünschte Verhalten unterlassen und wurde (infolge (25)) nicht von der Umwelt  $u$  belohnt, sprich sie ist nicht in den Besitz von Käse gekommen. Zum Zeitpunkt  $k=1$  hat die Maus das gewünschte Verhalten ausgeführt, allerdings wurde sie dafür von der Umwelt dennoch nicht belohnt.<sup>17</sup> Der Käse wurde von der Umwelt vom Feld 3

<sup>16</sup>Eine Nicht-Bestrafung und Belohnung von non-konformen Verhalten ist damit ausgeschlossen. Und da wir grundlegend annehmen, dass es nicht möglich ist, ein und dasselbe Verhalten zugleich zu belohnen und zu bestrafen,  $\neg \mathbf{M}(G(u, m) \wedge P(u, m))$ , ist als Reaktion auf ein non-konformes Verhalten entweder eine Bestrafung und Nicht-Belohnung oder eine Nicht-Bestrafung und Nicht-Belohnung möglich.

<sup>17</sup>Was in diesem abgewandelten dynamischen Szenario also nicht mehr gilt, ist der Sachverhalt, dass die Maus bei Ausführung des entsprechenden Verhaltens in den Besitz von Käse kommt. Es gilt also nicht wie im ersten Szenario neben  $N(K(m) \rightarrow V_{F_{2m}O}(m))$  auch noch  $N(V_{F_{2m}O}(m) \rightarrow K(m))$ . Letzteres ist logisch äquivalent mit  $\neg \mathbf{M}(V_{F_{2,0}O}(m) \wedge \neg K(m))$ , allerdings ist  $V_{F_{2,0}O}(m) \wedge \neg K(m)$  zum Zeitpunkt  $k=1$  der Fall, sodass  $\mathbf{M}(V_{F_{2,0}O}(m) \wedge \neg K(m))$  gilt, was logisch äquivalent mit  $\neg N(V_{F_{2,0}O}(m) \rightarrow K(m))$  ist.

wegplatziert.<sup>18</sup> Zum Zeitpunkt  $k=2$  hat die Maus das gewünschte Verhalten nochmals ausgeführt und wurde von der Umwelt belohnt, kam also in den Besitz von Käse. Insgesamt ergibt sich aus der obigen Beschreibung folgende symbolische Übersicht:  $k=0: \neg V_{F2,O}(m) \wedge \neg G(u, m)$ ;  $k=1: V_{F2,O}(m) \wedge \neg G(u, m)$ ;  $k=2: V_{F2,O}(m) \wedge G(u, m)$ . Entsprechend der Modellierung von Zeitpunkten als mögliche Welten, liegt hier zu keinem Zeitpunkt (in keiner möglichen Welt) der Sachverhalt vor, dass die Maus das gewünschte Verhalten unterlassen hat und von der Umwelt belohnt wurde:  $\neg \mathbf{M}(\neg V_{F2,O}(m) \wedge G(u, m))$ . Dies wiederum ist (entsprechend den Ausführungen in Abschnitt 3) logisch äquivalent mit  $\mathbf{N}(G(u, m) \rightarrow V_{F2,O}(m))$ , womit wir (und die Maus) das Abzuleitende abgeleitet hätte:<sup>19</sup> den Sachverhalt, dass ein bestimmtes Verhalten eine notwendige Bedingung für Belohnung darstellt:  $\mathbf{N}(G(u, m) \rightarrow V_{F2,O}(m))$ .

Eine konsequente Reaktion, hier eine Belohnung,<sup>20</sup> ist im Falle konformen Verhaltens (wie in 23) also nicht notwendig. Allerdings heißt das nicht, dass jegliche Reaktion möglich wäre. Von der Umwelt gewünschtes Verhalten (also unter der Bedingung von  $\mathbf{O}^u(V(m))$ ) darf nämlich weder konsequent bestraft noch konsequent nicht belohnt werden [19]:

$$(27) \quad \neg \mathbf{N}(V(m) \rightarrow P(u, m))$$

$$(28) \quad \neg \mathbf{N}(G(u, m) \rightarrow \neg V(m))$$

Wäre das der Fall, so würde aufgrund des zu keinem Zeitpunkt auftretenden  $V(m) \wedge \neg P(u, m)$  und  $V(m) \wedge G(u, m)$  sowohl  $\neg \mathbf{M}(V(m) \wedge \neg P(u, m))$  als auch  $\neg \mathbf{M}(V(m) \wedge G(u, m))$  gelten und damit  $\mathbf{N}(V(u) \rightarrow P(u, m))$  und  $\mathbf{N}(G(u, m) \rightarrow \neg V(m))$  folgen. Zusammen mit  $\mathbf{N}(\neg V(m) \rightarrow P(u, m))$ ,  $\mathbf{N}(G(u, m) \rightarrow V(m))$  und dem Sachverhalt, dass der Agent sowohl Belohnungen als auch Nicht-Bestrafungen intendiert, hätte dies jeweils einen Widerspruch zur Folge, sodass die Umwelt hinsichtlich ihrer Reaktion auf von ihr gewünschtes Verhalten des Agenten nicht vollkommen frei ist: Sie muss von ihr gewünschtes Verhalten entweder konsequent belohnen oder nicht-bestrafen oder aber dahingehend zumindest konsequent inkonsequent sein, indem sie das gewünschte Verhalten mal belohnt und nicht belohnt, mal bestraft und nicht bestraft.

## 6. Belohnung und Bestrafung

Entscheidend ist, dass der kognitive Agent so nicht wissen muss, was die Umwelt (als lehrender kognitiver Agent) will oder nicht will. Das Erlernen des seitens des lernenden Agenten zu Wollenden oder nicht zu Wollenden (das dem seitens der Umwelt Gewolltem oder Nicht-Gewolltem entspricht) erfolgt indirekt über Bestrafung und Belohnung. Was jedoch eine Belohnung oder Bestrafung ist, hängt davon ab, ob der Reiz vom jeweiligen Agenten als positiv oder negativ klassifiziert wird. Eine Belohnung liegt bei Eintreten eines positiven Reizes,  $R_P(u, m)$ , vor, eine Bestrafung hingegen bei Eintreten eines negativen Reizes,  $R_N(u, m)$ :<sup>21</sup>

$$(29) \quad G(u, m) \leftrightarrow R_P(u, m)$$

$$(30) \quad P(u, m) \leftrightarrow R_N(u, m)$$

<sup>18</sup>Wie wir in der ersten Anmerkung angeführt haben, gehen wir davon aus, dass die Umwelt selbst ein kognitiver Agent sein kann (aber nicht sein muss).

<sup>19</sup>Aufgrund des erschlossenen Wissens um  $\mathbf{N}(G(u, m) \rightarrow V_{F2,O}(m))$  kann sie mithilfe des deontischen modus ponens und  $\mathbf{O}^m(G(u, m))$ , ihrem Wunsch nach Belohnung, auf  $\mathbf{O}^m(V_{F2,O}(m))$  schließen: dem Sollen das entsprechende Verhalten auszuführen. Wie zu sehen, bedarf die Maus tatsächlich weder des Wissens um (24) und (25).

<sup>20</sup>Dasselbe gilt auch für Nicht-Bestrafungen als Reaktion auf konformes Verhalten. Vgl. hierzu ausführlich [19] in diesem Band.

<sup>21</sup> $R_P(u, m)$  ist zu lesen als „u setzt m einem positiven Reiz aus“ und  $R_N(u, m)$  als „u setzt m einem negativen Reiz aus“.



Entsprechend stellt eine Nicht-Belohnung das Ausbleiben eines positiven Reizes,  $\neg R_P(u, m)$ , eine Nicht-Bestrafung das Ausbleiben eines negativen Reizes,  $\neg R_N(u, m)$ , dar:

$$(31) \quad \neg G(u, m) \leftrightarrow \neg R_P(u, m)$$

$$(32) \quad \neg P(u, m) \leftrightarrow \neg R_N(u, m)$$

Was allerdings ein positiver oder ein negativer Reiz ist, hängt wiederum vom jeweiligen lernenden Agenten ab, was sich dahingehend formal beschreiben lässt, als dass ein eingetretener vom lernenden Agenten gewollter Reiz ein positiver Reiz ist und ein ausgebliebener vom lernenden Agenten gewollter Reiz ein negativer Reiz ist.

$$(33) \quad \mathbf{O}^m(R(u, m)) \wedge R(u, m) \rightarrow R_P(u, m)$$

$$(34) \quad \mathbf{O}^m(R(u, m)) \wedge \neg R(u, m) \rightarrow R_N(u, m)$$

Analog gilt hinsichtlich nicht-gewollter Reize, dass ein ausgebliebener vom lernenden Agenten nicht-gewollter Reiz ein positiver Reiz ist und ein eingetretener vom lernenden Agenten nicht-gewollter Reiz ein negativer Reiz ist:

$$(35) \quad \mathbf{O}^m(\neg R(u, m)) \wedge \neg R(u, m) \rightarrow R_P(u, m)$$

$$(36) \quad \mathbf{O}^m(\neg R(u, m)) \wedge R(u, m) \rightarrow R_N(u, m)$$

Nicht-negative und nicht-positive Reize sind hingegen das Resultat des Eintretens oder Ausbleibens von Reizen, deren Wollen oder Nicht-Wollen nicht zutrifft.<sup>22</sup>

$$(37) \quad \neg \mathbf{O}^m(R(u, m)) \wedge R(u, m) \rightarrow \neg R_P(u, m)$$

$$(38) \quad \neg \mathbf{O}^m(R(u, m)) \wedge \neg R(u, m) \rightarrow \neg R_N(u, m)$$

$$(39) \quad \neg \mathbf{O}^m(\neg R(u, m)) \wedge \neg R(u, m) \rightarrow \neg R_P(u, m)$$

$$(40) \quad \neg \mathbf{O}^m(\neg R(u, m)) \wedge R(u, m) \rightarrow \neg R_N(u, m)$$

Betrachtet man die Relation von  $R(u, m)$  und  $\neg R(u, m)$  innerhalb der Formeln und bedenkt, dass es sich hierbei um Prädikatsvariablen handelt, so genügen die Formeln (33), (34), (37) und (38).

## 7. Ausblick

Wir haben gezeigt, dass situationsabhängige Verhaltensvorschriften für kognitive Agenten aufbauend auf bestimmten Schlussverfahren aus einer kleinen Menge an basalen Präskriptionen und Intentionen sowie der Interaktion mit der Umwelt formallogisch deduziert werden können. Für eine technische Anwendung muss nun, beispielsweise für das Maus-Labyrinth-Beispiel, eine solche minimale Menge gefunden und anhand von Simulationen gezeigt werden, dass der Agent das nötige Wissen tatsächlich ermitteln kann.

## Literatur

[1] BISCHOF, N.: *Psychologie, ein Grundkurs für Anspruchsvolle*. Kohlhammer, Stuttgart 2014.

---

<sup>22</sup>Spätestens an dieser Stelle stößt die (deutsche) natürliche Sprache an ihre Grenzen: Natürlichsprachlich wird das Nicht-Wollen von p als das Wollen von nicht-p (bzw. das Wollen, dass p nicht zutrifft) verstanden. Ein solches Wollen bzw. Nicht-Wollen entspricht einem  $\mathbf{O}(\neg p)$ , wovon  $\neg \mathbf{O}(p)$  zu unterscheiden ist: In diesem Fall trifft nicht zu, dass p gewollt ist.

- [2] KLIMCZAK, P.; WOLFF, M.; LINDEMANN, J.; PETERSEN, C.; RÖMER, R.; ZOGLAUER, T.: Die kognitive Heizung. In: *Elektronische Sprachsignalverarbeitung*, 24. 2014, S. 89 – 96.
- [3] WATKINS, C., DAYAN, P.: Q-learning. In: *Machine learning*, 8 (3-4). 1992, S. 279 – 292
- [4] MNH, V., ET AL.: Human-level control through deep reinforcement learning. In: *Nature* 518. 2014, S. 529 – 533.
- [5] HUME, D.: *A Treatise of Human Nature*. Oxford University Press, Oxford 1978.
- [6] KUTSCHERA, F.: Das Humesche Gesetz. In: *Grazer Philosophische Studien*, 4 (1). 1977, S. 1 – 4.
- [7] STUHLMANN-LAEISZ, R.: *Das Sein-Sollen-Problem. Eine modallogische Studie*. Frommann, Stuttgart 1983.
- [8] DÖRNER, D.: *Bauplan für eine Seele*. Rowohlt, Reinbek 1999.
- [9] SCHMIDT, I., RÖMER, R., WOLFF, M.: Semantische Repräsentation strukturierter Objekte im Hilbertraum und deren Verarbeitung nach quantenmechanischen Formalismen. In: *Elektronische Sprachsignalverarbeitung*, 27. 2017. (in diesem Band)
- [10] GRÄF, D., GROSSMANN, S., KLIMCZAK, P., KRAH, H., WAGNER, M.: *Filmsemiotik. Eine Einführung in die Analyse audiovisueller Formate*. Schüren, Marburg 2011.
- [11] HUGHES, G. E., CRESSWELL, M. J.: *A New Introduction to Modal Logic*. Routledge, London 1996.
- [12] ZOGLAUER, T.: *Einführung in die formale Logik für Philosophen*. Göttingen, Vandenhoeck & Ruprecht, 2008.
- [13] STEGMÜLLER, W.: *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Band I: Erklärung – Begründung – Kausalität*. Springer, Berlin 1983.
- [14] KLIMCZAK, P.: *Formale Subtextanalyse. Kalkülisierung von Narration und Interpretation*. Mentis, Münster, 2016.
- [15] ZOGLAUER, T.: *Normenkonflikte. Zur Logik und Rationalität ethischen Argumentierens*. Frommann-Holzboog, Stuttgart 1998.
- [16] MORSCHER, E.: *Normenlogik. Grundlagen – Systeme – Anwendungen*. Mentis, Paderborn 2012.
- [17] BAYER, K.: *Argument und Argumentation. Logische Grundlagen der Argumentationsanalyse*. Vandenhoeck & Ruprecht, Göttingen 2007.
- [18] NAEISS, A.: *Kommunikation und Argumentation. Eine Einführung in die angewandte Semantik*. Scriptor, Kronberg 1975.
- [19] KLIMCZAK, P., WIRSCHING, G.: Formale Analysen des ABC-Schemas und des operanten Konditionierens. In: *Elektronische Sprachsignalverarbeitung*, 27. 2017. (in diesem Band)
- [20] JOERDEN, J. C.: *Logik im Recht. Grundlagen und Anwendungsbeispiele*. Springer, Berlin 2005.
- [21] ANDERSON, J. R.: *Learning and memory: An integrated approach*. Wiley, New York 2000.
- [22] HAGMAYER, Y.: Logik in der Psychologie. Warum Menschen nicht gemäß den Gesetzen der Logik schlussfolgern. In: KLIMCZAK, P., ZOGLAUER, T.: *Logik in den Wissenschaften*. Mentis, Münster 2017. (im Erscheinen)
- [23] DA SILVA NEVES, R., BONNEFON, J.-F., RAUFASTE, E.: An empirical test of patterns of nonmonotonic inference. In: *Annals of Mathematics and Artificial Intelligence*, 34. 2002, S. 107 –130.
- [24] DE HOUWER, J.: The propositional approach to associative learning as an alternative for association formation models. In: *Learning & Behavior*, 37. 2009, S. 1 – 20.
- [25] RESCORLA, R. A.: Pavlovian Conditioning. It's Not What You Think It Is. In: *American Psychologist*, 43. 1988, S. 151 – 160.