

ENTSCHEIDUNGSTHEORETISCHE MODELLIERUNG DER KONSUMMATORISCHEN ENDHANDLUNG – VERGLEICH VON KLASSISCHEN UND QUANTENMECHANISCHEN ANSÄTZEN

Ronald Römer, Peter beim Graben und Matthias Wolff

BTU Cottbus-Senftenberg

ronald.roemer@b-tu.de

Kurzfassung: Ein wesentlicher Bestandteil kognitiver – und damit auch sprachverarbeitender – Agentensysteme ist die Verhaltenssteuerung. Eine ihrer wichtigsten Aufgaben besteht darin, Entscheidungsprobleme, die in der Auseinandersetzung mit der physischen Umgebung eines Agenten auftreten, rational zu lösen. Zunächst diskutieren wir die klassischen, nutzenbasierten Entscheidungsmodelle für sequentielle- und episodische Umgebungen am Beispiel des Mouse-Maze-Problems. Zu letzteren gehören die sogenannten konsummatorischen¹ Endhandlungen, sie stehen im Mittelpunkt dieser Arbeit. Weil sich Entscheidungen bei Endhandlungen final auf die Lebensdauer des Agenten auswirken können, dürfen Fehlleistungen und Täuschungen nur äußerst selten auftreten. Aus diesem Grund stellen wir der rationalen Methode die psychologische Methode entgegen, die zusätzlich auf der Einführung von semantischen Kategorien beruht. Die dafür benötigten semantischen Repräsentationen modellieren wir nach einem quantenmechanischen Formalismus. Daraus leitet sich die zentrale Fragestellung dieser Arbeit ab: Bietet die quantenmechanische Perspektive einen Mehrwert bei der Modellierung der konsummatorischen Endhandlung und lässt sich dieser Formalismus in den Perzeptions-Aktionszyklus (PAC) eines kognitiven Agenten einbetten.

1 Motivation und Einführung

In [2] haben wir einen Ausblick auf einen Bauplan für technische kognitive Systeme nach den Vorgaben einer systemtheoretisch geprägten Psychologie [1] gegeben. Diesen Bauplan haben wir in Bezug auf Funktion und Struktur kognitiver Systeme weiter konkretisiert und nach dem Vorbild des Mouse-Maze-Problems ein Minimalsystem vorgestellt, das als Experimentiersystem in Lehre und Forschung eingesetzt wird [3]. Zu den Forschungsschwerpunkten gehören Wissensrepräsentation und Problemlösung sowie Kommunikation und Lernen [4]. Die vorliegende Arbeit beschäftigt sich schwerpunktmäßig mit der Modellierung von konsummatorischen Endhandlungen, welche sich an die Problemlösung (z.B. Auffindung von Zielobjekten) anschließen. Bei dem hier betrachteten Mouse-Maze-Problem ist das Zielobjekt ein Stück Käse, welches in den Zuständen „frisch“ oder „verdorben“ vorliegen kann. Diese Zustände kann die Maus über die Farbwahrnehmungen „blau“ oder „weiß“ unterscheiden und sich anschließend für eine der beiden Endhandlungen „Aufnahme“ bzw. „Verweigerung“ entscheiden. Nach Ausführung der konsummatorischen Endhandlung „Aufnahme“ wird der Antrieb zur Bedürfnisbefriedigung gelöscht und die Verhaltenssteuerung kann in ihren homöostatischen Gleichgewichtszustand überführt werden.

¹Konsummatorisch sollte nicht mit konsumatorisch verwechselt werden. Ersteres orientiert sich am lateinischen Ursprung „consummare“ und meint in der Psychologie den vollendenden Aspekt einer Handlungskette [1].

2 Problemstellung und Zielsetzung

Bei Entscheidungsproblemen ist es für einen Agenten nicht ausreichend, lediglich Ziele, Zustände und Aktionen zu definieren. Um rationales Verhalten nachzubilden, ist es wünschenswert den vorzuziehenden Weltzuständen einen größeren Nutzen zuzuordnen. Häufig basieren solche nutzenbasierte Modelle auf Markov-Entscheidungsprozessen (MDP, POMDP), wobei wir hier zwei Anwendungen unterscheiden. Bei der Zielfindung liegt der Schwerpunkt auf dem Erlernen einer Strategie, mit der der Agent eine sequentielle Folge von Entscheidungen realisieren kann. Diese Entscheidungsprozesse lassen sich problemlos in den PAC (Abbildung 1) kognitiver Systeme integrieren [5].

Bei der konsummatorischen Endhandlung liegt der Anwendungsschwerpunkt auf der semantischen Bewertung des vorgefundenen Zielobjekts. Da wir die dafür nötigen semantischen Repräsentationen nach einem quantenmechanischen Formalismus modellieren, müssen wir nach einer Brücke zwischen den beiden Modellierungsansätzen suchen. Damit ist die Zielstellung

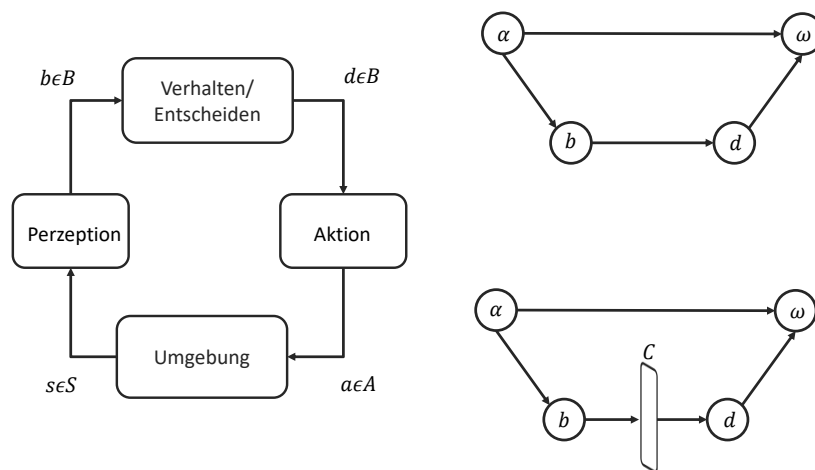


Abbildung 1 - Modelle für Entscheidungsprozesse. Links - Modell für das sequentielle Entscheidungsproblem zur Auffindung von Umgebungszuständen: Sensorik und Aktuatorik werden durch bedingte Wahrscheinlichkeiten $p(b|s)$ bzw. $p(a|d)$ modelliert. Rechts (oben) - konsummatorische Endhandlung nach [6]: Beobachtungen b von Objektzuständen α führen zu Entscheidungen d . Diese ziehen Folgezustände ω nach sich. Rechts (unten): Kategorisierung als mentales Werkzeug eines kognitiven Agenten.

der vorliegenden Arbeit vorgegeben. Wir wollen untersuchen, ob sich der quantenmechanische Formalismus bei der Modellierung von konsummatorischen Endhandlungen gewinnbringend einsetzen lässt und ob sich dieser Formalismus in den PAC integrieren und in unserem Minimalsystem anwenden lässt.

2.1 Lösungsansätze und Methodik

Methodisch folgen wir dem demiurgischen Prinzip und konstruieren den Agenten als ein finales System, dessen Struktur eine „adaptive“ Funktion zu erfüllen hat, mit der Schäden vermieden werden sollen [6]. Der wissenschaftstheoretische Orientierungsrahmen bezieht sich demnach auf die „konstruktive“ Methode. In der vorliegenden Arbeit untersuchen wir drei Ansätze: 1) entscheidungstheoretischer Ansatz auf der Grundlage von nutzenbasierten Agenten, 2) klassischer Markov-Entscheidungsprozess und 3) quantenmechanischer Entscheidungsprozess. Der erste Ansatz entspricht der Sicht eines äußeren Beobachters. Er hat keine Kenntnis über innere

Vorgänge des Agenten. Allerdings lassen sich die möglichen Ereignisse, Beobachtungen und Entscheidungen statistisch durch eine Verbundwahrscheinlichkeit erfassen. Wie bei den Bayes-netzen kann nun unter Berücksichtigung von bedingten stochastischen Unabhängigkeiten die Verbundverteilung so faktorisiert werden, dass eine Kausalkette erzeugt wird.

Mit der Verkettung können wir zu einer inneren Sichtweise des Agenten übergehen, so dass der Agent mit der Kausalkette experimentieren kann. Diese Idee wird von den anderen beiden Ansätzen verfolgt. Zunächst fügt der Agent bei der experimentellen Analyse zwischen Wahrnehmung B und Endhandlung D eine interne Kategorisierungsstufe C in die Markovkette ein (Abbildung 1, rechts unten) und stuft dabei das Zielobjekt in die semantischen Kategorien genießbar oder ungenießbar ein. Damit können von der Wahrnehmung B über die Kategorisierung C bis zur Entscheidung D verschiedene Wege durch den entstandenen Markov-Graphen genommen werden (Abbildung 2, rechts unten). Die Struktur eines solchen Experiments entspricht der Struktur des Doppelspaltexperiments in der Physik [7].

Im klassischen Fall – hier eine Markovkette – wird immer nur ein einziger Pfad durch den Markov-Graphen durchlaufen (siehe Generatormodell nach Abbildung 2, rechts oben). Dies entspricht der experimentellen Voreinstellung, wonach der tatsächliche Weg beobachtet werden kann (Analogie: Licht als Partikel). Komplementär dazu ist die experimentelle Voreinstellung, nach der alle Wege gleichzeitig genommen werden (Analogie: Licht als Welle). Experimente mit menschlichen Probanden haben gezeigt, dass das quantenmechanische Modell für die tatsächlich gemessenen Wahrscheinlichkeiten möglicher Endhandlungen die genaueren Vorhersagen liefern kann [7]. Dies ist auf einen zusätzlichen Interferenzterm zurückzuführen, der bei der Überlagerung von Wahrscheinlichkeitswellen entsteht.

2.2 Zur Interpretation des Doppelspaltexperiments

Im Mittelpunkt steht die Diskussion um das Auftreten der Superposition. In der makroskopischen Welt der Objekte wird sie bekanntlich überhaupt nicht beobachtet, als mathematisches Konzept ist sie dagegen unproblematisch. Dies führte zu zwei einander entgegengesetzten Auffassungen:

- entweder die Superposition wird ernstgenommen, wobei die Beschreibung des Phänomens mathematisch mit der Wellenfunktion erfolgt (Messungen am Objekt führen zur Aufspaltung der Realität: Viele-Welten-Interpretation), oder
- die Superposition ist obsolet (deBroglie-Bohm: eine sogenannte Führungswelle leitet das Partikel zum Interferenzschirm).

Eine experimentelle Variante des Doppelspaltexperiments ist das Mach-Zehnder-Interferometer. Bei diesem Aufbau (siehe Abbildung 2, links) wird die Intensität des Lichts soweit reduziert, dass immer nur ein Photon (d.h. ein Partikel) die Quelle verlässt. Detektoren an den beiden Ausgängen des Interferenzaufbaus stellen das Eintreffen des Teilchens fest. Damit beschränkt sich die Beschreibung des Beobachters auf *logische Aussagen* an den Detektoren und führt auf die Kopenhagener Interpretation nach Niels Bohr. Danach ist die Wellenfunktion nur ein mathematisches Hilfsmittel, mit der Wahrscheinlichkeiten berechnet werden können. Die Ausbreitung einer Welle in Zeit und Raum entspricht demnach nicht der physikalischen Realität.

Die wichtigste Konsequenz dieser Interpretation ist das Phänomen der Komplementarität. Bei der experimentellen Voreinstellung *ohne Interferenzspiegel* m_1 führt die Beobachtung an den Detektoren zur sicheren Aussage, welchen der beiden Wege das Partikel genommen hat. Bei der Voreinstellung *mit Interferenzspiegel* m_1 hängen die Aussagen davon ab, wie die Interferenzspiegel m_0 und m_1 arrangiert wurden. Bei vollständiger destruktiver Interferenz an einem

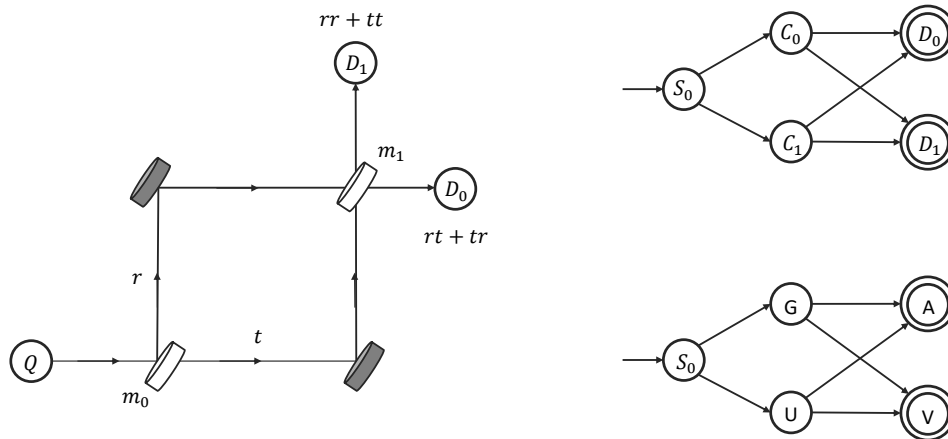


Abbildung 2 - Mach-Zehnder-Interferometer. Links - Der Aufbau umfasst eine Photonenquelle Q , zwei halbdurchlässige Spiegel (m_0, m_1), zwei reflektierende Spiegel sowie zwei Detektoren (D_0, D_1). Am rechten Detektor wird konstruktive Interferenz festgestellt, da sowohl der obere als auch der untere Pfad je eine Reflektion mit gleicher Phasenverschiebung aufweisen. Am oberen Detektor ist der obere Pfad durch zwei Reflektionen und der untere Pfad durch zwei Durchtritte gekennzeichnet. Diese Überlagerung führt zu destruktiver Interferenz. Rechts - Äquivalente Wegediagramme für das Interferometers (oben) und das Kategorisierungs-Entscheidungs-Experiment (unten).

Detektor und vollständiger konstruktiver Interferenz am anderen Detektor führt die Beobachtung zu der sicheren Aussage, welcher Detektor das Partikel empfängt. Für diese beiden experimentellen Voreinstellungen liefert das Experiment nun binäre Aussagen zu den Komplementärgrößen Weginformation und Interferenz, welche dadurch einer Kodierung zugänglich sind. Diese Aussagen können jedoch immer nur für einen konkreten experimentellen Aufbau getroffen werden, nicht für beide gleichzeitig. D.h. die Informationskapazität dieses Systems (ein Bit) genügt nicht, um beide Aussagen gleichzeitig zu kodieren [8].

Wird nun der experimentelle Aufbau so ausgeführt, dass die destruktive bzw. konstruktive Interferenz an den Ausgangsdetektoren durch Parametrierung (Grad der Durchlässigkeit der Spiegel, einstellbare Weglängen) gestört werden kann, dann führt diese Unsicherheit dazu, dass für beide Komplementärgrößen nur noch Wahrscheinlichkeiten angegeben werden können, was dann durch die Verwendung des Qubits zum Ausdruck kommt.

3 Grundlagen der Entscheidungstheorie

Charakteristisch für Markovprozesse ist die Berücksichtigung unsicherer oder unvollständiger Information. Für den Agenten bedeutet das, dass er Prioritäten in Bezug auf die herbeizuführenden Ergebnisse setzen muss. Folglich müssen in der Entscheidungstheorie Prioritäten, die durch den Nutzen ausgedrückt werden, mit Wahrscheinlichkeiten kombiniert werden.

Strategie und Nutzen. Für die Berechnung des Nutzens bzw. einer Nutzenfunktion gehen wir von zwei Mengendefinitionen aus. Zunächst definieren wir die Menge der Ereignisse $E = \{e_1, e_2, \dots, e_N\}$, für die der Nutzen $U(e_i)$ ermittelt werden soll. Mit der zweiten Menge $D = \{d_1, d_2, \dots, d_M\}$ werden die möglichen Entscheidungen festgelegt. Wir definieren den Nutzen $U(e_i)$ eines Ereignisses e_i als Erwartungswert des bedingten Nutzens $U(e_i|d_k)$, d.h. dem Nutzen

des Ereignisses e_i bei gegebener Entscheidung d_k .

$$U(e_i) = \sum_{k=1}^M p(d_k) \cdot U(e_i|d_k)$$

Dabei gilt für die Indizes $i \in \{1..N\}$ und $k \in \{1..M\}$. Um den mittleren Nutzen $U(E)$ zu erhalten, bestimmen wir den Erwartungswert über alle möglichen Ereignisse e_i :

$$\begin{aligned} U(E) = E\{U(e)\} &= \sum_{i=1}^N p(e_i) \cdot U(e_i) \\ &= \sum_{i=1}^N p(e_i) \sum_{k=1}^M p(d_k) \cdot U(e_i|d_k) \end{aligned}$$

Als Nutzenfunktion werden häufig $U(e_i|d_k) = \delta_{i,k}/p(e_i)$ oder $U(e_i|d_k) = \delta_{i,k} \cdot \log(1/p(e_i))$ verwendet. Die Verteilung $p(d_k)$ wird Strategie genannt. Bei der Anwendung des *Prinzips des maximalen erwarteten Nutzens* werden mit den getroffenen Entscheidungen diejenigen Ereignisse ausgewählt, welche den höchsten erwarteten Nutzen einbringen

$$d^* = \arg \max_{p(d)} [U(E)].$$

Prognostische und diagnostische Sichtweise. Bei der Kombination mit Nutzenfunktionen spielen zusätzlich bedingte Wahrscheinlichkeitsverteilungen eine wichtige Rolle. In Bezug auf den rationalen Agenten sind die transitiven Verteilungen $p(b|e)$ für die Sensorik und $p(a|d)$ für die Aktuatorik zwei Beispiele. Dabei werden mit der Menge $B = \{b_1, b_2, \dots, b_N\}$ die Beobachtungen an den Sensoren und mit $A = \{a_1, a_2, \dots, a_M\}$ die Menge der physischen Aktionen der Aktuatorik formal erfasst. Bedingte Verteilungen erhält man aus den Verbundverteilungen, welche Zufallsvariablen zueinander in Beziehung setzen. Welche der Zufallsvariablen hierbei Ursache oder Wirkung ist, wird von dieser Relation jedoch nicht erfasst. Das wird daran deutlich, dass sich Verbundverteilungen auf symmetrische Weise faktorisieren lassen. Für die Sensorik gilt beispielsweise $p(e, b) = p(b|e) \cdot p(e) = p(e|b) \cdot p(b)$.

Für eine kausale Betrachtungsweise geordnet nach Ursache (Ereignis E) und Wirkung (Beobachtung B) sind die beiden bedingten Wahrscheinlichkeiten notwendig. Fragt man nach den Folgen $b \in B$ der Ereignisse $e \in E$, dann gilt die *prognostische Sichtweise* mit $p(b) = \sum_e p(b|e)p(e)$, was einer Vorhersage in Transitionsrichtung entspricht. Fragt man dagegen nach den Ursachen $e \in E$ bei gegebenen Wirkungen $b \in B$, so gilt für die *diagnostische Sichtweise* gemäß des Bayes'schen Satzes

$$p(e|b) = \frac{p(b|e)p(e)}{p(b)}.$$

Für die formale Beschreibung der nachfolgenden entscheidungstheoretischen Modelle werden beide Sichtweisen benötigt.

4 Entscheidungstheoretische Modelle nutzenbasierter Agenten

Bei der Übertragung des Entscheidungsproblems auf den PAC spielen die Beobachtungen an den entsprechenden Sensoren eine wichtige Rolle. Für die beiden Entscheidungsprobleme (Zielfindung, Zielbewertung) gehen wir zunächst davon aus, dass der Agent vollständigen Zugriff auf den Zustand der Umgebung hat und dass die Entscheidungen aus D in einer 1:1 Beziehung zu den Aktionen aus A stehen. In diesem Fall gilt für die Beobachtung der möglichen Umgebungs-

bzw. Objektzustände $b = s \in S$ und für die Aktionen $a = d \in D$. Folglich tritt sowohl an der Zustandssensorik als auch an der Aktuatorik keine Streuung auf. Beim sequentiellen Entscheidungsproblem kommt es nun darauf an, den erwarteten Nutzen $U(S, D)$ zu maximieren und damit den vorzuziehenden Paaren (s, d) die notwendige Priorität zu verleihen. Dagegen liegt bei der konsummatorischen Endhandlung der Fokus auf der Sensorik, so dass bei gegebener Beobachtung der bedingte mittlere Nutzen $U(S|B)$ maximiert werden muss.

4.1 Sequentielles Markov-Entscheidungsproblem

Bei diesem Modell optimiert der nutzenbasierte Agent die Wirksamkeit seiner Entscheidungen $d \in D$ auf den Umgebungszustand $s \in S$. Das Modell der Umgebung beschränkt sich auf die transitive Verteilung $p(s'|s)$. Um nun ein bestimmtes Ziel zu erreichen, muss der Agent die Umgebung wirksam manipulieren können. Dafür wird dann ein gesteuerter stochastischer Markovprozess $p(s'|s, d)$ benötigt. Für das Mouse-Maze-Problem ist der Zustand durch die x, y -Koordinaten des Labyrinths bestimmt, die Entscheidungen des Agenten beziehen sich hier auf die Bewegungsrichtungen. Der entscheidungstheoretische Ansatz zur Berechnung des mittleren erwarteten Nutzens lautet:

$$\begin{aligned} U(S, D) &= E\{U(s, d)\} = \sum_s \sum_d p(s, d) \cdot U(s, d) \\ &= \sum_s p(s) \sum_d p(d|s) \cdot U(s, d). \end{aligned}$$

Die optimale Strategie, mit der man den mittleren erwarteten Nutzen maximieren kann, erhält man mit

$$\pi^* = p(d^*|s) = \arg \max_{p(d|s)} [U(S, D)].$$

Im Rahmen des sequentiellen Markov-Entscheidungsproblems wird zur Bestimmung der konkreten Nutzenfunktion $U(s, d)$ die Bellmangleichung verwendet [9]. Darin wird iterativ für jeden Zustand $s \in S$ der Informationswert $V^\pi(s)$ bei gegebener Strategie berechnet

$$V^\pi(s) = \sum_d p(d|s) \sum_{s'} p(s'|d, s) [r(s, d, s') + V^\pi(s')].$$

Neben dem stochastischen Modell des gesteuerten Markov-Entscheidungsprozesses wird für jede Transition eine aus der Umgebung stammende lokale Belohnung $r(s, d, s')$ angenommen. Zur Bestimmung des Informationswerts muss nun für jede Entscheidung d über alle möglichen Nachfolgezustände s' von s summiert werden. Neben dem Zusammenhang zum mittleren erwarteten Nutzen

$$U^\pi(S, D) = \sum_s p(s) \sum_d p(d|s) \sum_{s'} p(s'|d, s) [r(s, d, s') + V^\pi(s')],$$

also $U^\pi(S, D) = \sum_s p(s) V^\pi(s)$, wird damit auch die Bestimmung der konkreten Nutzenfunktion bei gegebener Strategie π verdeutlicht:

$$U^\pi(s, d) = \sum_{s'} p(s'|d, s) [r(s, d, s') + V^\pi(s')].$$

In der Fachliteratur zum Reinforcement-Learning wird die Nutzenfunktion $U(s, d)$ auch als Q-Funktion bezeichnet und diese als Namensträger für das Q-Learning verwendet. Beim Mouse-Maze-Problem setzen wir dieses Lernverfahren für das Auffinden von Zielobjekten ein [3].

4.2 Entscheidungsproblem der konsummatorische Endhandlung

Entscheidungen für abschließende Endhandlungen können ebenfalls nach dem Prinzip des maximalen Nutzens modelliert werden. Dabei erfolgt zusätzlich eine Kategorisierung der gefundenen Zielobjekte. Kategorien sind nach [6] als Werkzeug des Wahrnehmungsapparats finaler Systeme zu betrachten. Die zugehörige Entscheidung, ob ein Zielobjekt überhaupt für die Befriedigung eines Bedürfnis geeignet ist, fällt damit bereits in den Bereich der Semantik.

Für die weitere Betrachtung müssen wir die bislang verwendeten Zufallsvariablen neu interpretieren. Die möglichen Ereignisse $e \in E$ sind hier Zielobjekte z.B. verschiedene Käseobjekte, deren Zustände durch $S = \{\text{frisch, verdorben}\}$ gegeben sind. In unserem Beispiel soll der Zugriff auf die Objektzustände direkt erfolgen, dies wird in Abbildung 1 mit den Notationen $s \in S$ (links) sowie α (rechts) verdeutlicht. Weiterhin sind Beobachtungen nun nicht mehr die logischen Variablen eines binären Käse-Klassifikators sondern binäre Farbwahrnehmungen aus $B = \{\text{blau, weiß}\}$. Die Menge der Entscheidungen ist durch $D = \{\text{Aufnahme, Verweigern}\}$ bestimmt. Für die Modellierung der abschließenden Endhandlung benötigen wir einen stochastischen Entscheidungsprozess, der die zweiwertigen Zufallsvariablen S, B und D umfasst. Die Modellgüte hängt dabei vor allem von der Genauigkeit der Beobachtungen an den Zustandssensoren ab. Folglich kommt es jetzt darauf an, den bedingten mittleren Nutzen $U(S|B)$ und damit den Nutzen der Sensorik bzw. der Beobachtungen zu maximieren.

Das gemeinsame Auftreten der möglichen Zustände, Beobachtungen und Entscheidungen wird mit der Verbundwahrscheinlichkeit $p(s, b, d)$ erfasst. Wie bei den Bayesnetzen kann unter Berücksichtigung der bedingten stochastischen Unabhängigkeit $p(s|b, d) = p(s|b)$ die Verbundverteilung so faktorisiert werden, dass die Kausalkette $p(s, b, d) = p(s|b)p(d|b)p(b)$ erzeugt wird. Ausgangspunkt der folgenden Sichtweisen ist die Beobachtung b . Die bedingte Wahrscheinlichkeit $p(s|b)$ entspricht der diagnostischen Sichtweise: Ausgehend von einer konkreten Beobachtung soll auf die Ursache rückgeschlossen werden. Dagegen entspricht die bedingte Wahrscheinlichkeit $p(d|b)$ der prognostischen Sichtweise: Ausgehend von der konkreten Beobachtung soll eine Entscheidung vorhergesagt werden. Damit lässt sich nun auch für diesen Fall ein nutzenbasiertes Modell für den PAC angeben. Für den bedingten Nutzen folgt bei gegebener Beobachtung

$$U(s|b) = \sum_d p(d|b) \cdot U(s|d).$$

Die Berechnung des Erwartungswerts des bedingten Nutzens führt dann auf

$$\begin{aligned} U(S|b) &= \sum_s p(s|b) \sum_d p(d|b) U(s|d) \\ &= \sum_s p(s|b) U(s|b). \end{aligned}$$

Da die Strategie $p(d)$ nun von der aktuellen Beobachtung abhängig ist, wird eine prognostische Sicht eingenommen, so dass $p(d)$ in $p(d|b)$ übergeht. Um den mittleren bedingten Nutzen der Sensorik zu bestimmen, muss jetzt noch über alle Beobachtungen gemittelt werden

$$\begin{aligned} U(S|B) &= \sum_b p(b) \sum_s p(s|b) U(s|b) \\ &= \sum_b p(b) \sum_s p(s|b) \sum_d p(d|b) U(s|d) \\ &= \sum_b \sum_s p(s, b) U(s|b). \end{aligned}$$

Bei dieser Mittelung wird nun die diagnostische Sicht eingenommen. D.h. mit der aktuellen Beobachtung geht $p(s)$ in die Verteilung $p(s|b)$ über, welche die möglichen Ursachen für die

Beobachtung erfasst. Die optimale Strategie ist diejenige bedingte Verteilung, welche den mittleren bedingten Nutzen maximiert

$$P(d^*|b) = \arg \max_{p(d|b)} [U(S|B)].$$

Im einfachsten Fall gilt für die Strategie das Kroneckerdelta

$$p(d|b) = \delta_{d,d^*} := \begin{cases} 1, & \text{für } d = d^*, \\ 0, & \text{sonst.} \end{cases}$$

Der mittlere bedingte Nutzen bestimmt sich dann mit

$$U(S|B) = \sum_b \sum_s p(s,b) U(s|d^*).$$

5 Quantenmechanisches Modell der konsummatorischen Enhandlung

Die bislang besprochenen Modelle beruhen auf der Verwendung einer Nutzenfunktion und geben Wege an, wie der Nutzen optimiert werden kann. Im Unterschied zu den *rationalen* Entscheidungstheorien, versuchen *psychologische* Entscheidungstheorien zu erklären, auf welche Weise Entscheidungen getroffen werden. Vergleichende Experimente haben gezeigt, dass die Vorhersagen der rationalen Modelle nicht immer mit dem tatsächlich Verhalten übereinstimmen [7]. Zur Erklärung kann man vorbringen, dass sich die psychologischen Modelle in einem Punkt ganz wesentlich von den rationalen Modellen unterscheiden. Sie verwenden nach [1] *Kategorien*, mit denen die einströmenden Sinnesdaten geordnet werden (siehe Abbildung 2, rechts unten). Das Einfügen von Kategorien erlaubt es dem Agenten, durch Experimentieren auf die Natur der Wahrnehmungsinhalte zu schließen. Wenn dies möglichst vorurteilsfrei geschehen soll, dann sind quantenmechanische Ansätze wegen ihrer komplementären Erklärungskraft besonders geeignet. Das lässt sich nach [10] recht anschaulich bei der Interpretation von Systemzuständen in dynamischen Systemen verdeutlichen.

Gemäß der klassischen Denkfigur nimmt ein solches System zu jedem Zeitpunkt einen bestimmten Zustand ein. Eine Systembewegung wird dann – bei festgelegten Anfangsbedingungen – als Trajektorie durch den Zustandsraum beschrieben. Dies gilt auch für stochastische Systeme, hier herrscht lediglich Unkenntnis darüber, welches Element aus der Menge der möglichen Trajektorien tatsächlich realisiert wird. Den Elementen dieser Menge müssen daher Wahrscheinlichkeiten zugeordnet werden. Der klassische Modellansatz ähnelt damit der physikalischen Vorstellung, wonach sich ein *Partikel* genau einen Weg durch den Zustandsraum des Systems bahnt.

Dagegen geht die quantenmechanische Denkfigur zunächst davon aus, dass zu jedem Zeitpunkt jeder Zustand einen gewissen Anteil an der Systembewegung hat. Diese Ansatz deckt sich mit der physikalischen Vorstellung von einer *Welle* die sich durch den Zustandsraum bewegt. Das besondere an dieser Welle ist, dass jedem Zustand zu jedem Zeitpunkt eine Wahrscheinlichkeit zugeordnet wird – man spricht daher von einer Wahrscheinlichkeitswelle. Mit dem Wellencharakter rückt nun auch das Phänomen der Superposition in den Fokus der physikalischen Realität. Der Wellencharakter bleibt jedoch nur solange erhalten, bis eine Messung erfolgt. Der damit verbundene Kollaps der Wahrscheinlichkeitswelle offenbart dann wieder den Teilchencharakter. Während die Systembewegung also durch die Auslenkung unbestimmter Zustände erfolgt, führt die Messung schließlich wieder auf einen bestimmten Zustand. Dieser Ansatz integriert folglich sowohl die Wellen- als auch den Teilchenvorstellung. Beide Vorstellungen

sind insbesondere bei der Modellierung von konsummatorischen Endhandlungen nützlich. Unbestimmte Zustände trifft man hier häufig in Form von Widersprüchen oder Mehrdeutigkeiten vor der Handlungsentscheidung an. Erst mit der Auflösung dieser Schwierigkeiten liegt dann wieder ein bestimmter Zustand vor.

Formalisierung. Wir gehen nach Abbildung 1 wieder von einem direktem Zugriff auf den Objektzustand aus. D.h. anstelle der Farbwahrnehmungen „weiß“ und „blau“ liegt – je nach vorgefundenem Käseobjekt – unmittelbar der Startzustand $S = \text{„frisch“}$ oder $S = \text{„verdorben“}$ vor. Nach Abbildung 2 folgen dann die Übergänge in die Kategoriezustände der Menge $C = \{\underline{\text{Ungenießbar}}, \underline{\text{Genießbar}}\}$ und in die Zustände aus der Menge $D = \{\underline{\text{Aufnahme}}, \underline{\text{Verweigerung}}\}$. Wir betrachten im folgenden nur die Entscheidung $d = A$, da aus $d = V$ keine weiteren Einsichten folgen. Vom Startzustand S zum Zielzustand A gibt es zwei mögliche Wege $S \rightarrow G \rightarrow A$ und $S \rightarrow U \rightarrow A$. Wir betrachten das Zustandekommen der Aussage am Detektor für $d = A$ nun gemäß der Teilchen- oder Wellenvorstellung und berechnen die Wahrscheinlichkeitsfunktion $|\Psi(x,t)|^2$ gemäß den beiden folgenden physikalischen Ansätzen.

5.1 Modell für Partikelbahnen

Die beiden Wege werden durch die Wahrscheinlichkeitsfunktionen $|\Psi_1(x,t)|^2$ und $|\Psi_2(x,t)|^2$ beschrieben. Bei der anschließenden Summation werden folglich Wahrscheinlichkeiten addiert

$$|\Psi(x,t)|^2 = |\Psi_1(x,t)|^2 + |\Psi_2(x,t)|^2,$$

wobei $|\Psi(x,t)|^2 = \Psi(x,t) \cdot \Psi^*(x,t)$ gilt.

5.2 Modell für Wellenausbreitung

Bei diesem Modell werden die beiden Amplituden der Wellenfunktion $\Psi_1(x,t)$ und $\Psi_2(x,t)$ addiert und erst nachfolgend die Wahrscheinlichkeit berechnet.

$$\begin{aligned} |\Psi(x,t)|^2 &= |\Psi_1(x,t) + \Psi_2(x,t)|^2 \\ &= |\Psi_1(x,t)|^2 + |\Psi_2(x,t)|^2 + \Psi_1(x,t) \cdot \Psi_2^*(x,t) + \Psi_2(x,t) \cdot \Psi_1^*(x,t) \end{aligned}$$

Jede der beiden Amplituden der Wellenfunktion nimmt dabei den Wert einer komplexen Zahl $z = |z|e^{j\phi}$ an. Mit der vereinfachten Notation $z_1 = |z_1| \cdot e^{j\alpha}$ und $z_2 = |z_2| \cdot e^{j\beta}$ für die *Wahrscheinlichkeitsamplituden* sowie $z_1 \cdot z_2^* = |z_1| \cdot |z_2| e^{j(\alpha-\beta)}$ folgt schließlich

$$\begin{aligned} |z|^2 &= |z_1|^2 + |z_2|^2 + z_1 \cdot z_2^* + z_2 \cdot z_1^* \\ &= |z_1|^2 + |z_2|^2 + 2 \cdot \text{Re}\{z_1 \cdot z_2^*\} = |z_1|^2 + |z_2|^2 + 2 \cdot \text{Re}\{z_2^* \cdot z_1\} \\ &= |z_1|^2 + |z_2|^2 + 2 \cdot |z_1| \cdot |z_2| \cdot \cos(\Theta), \end{aligned}$$

wobei

$$\cos(\theta) = \frac{\text{Re}\{z_1 \cdot z_2^*\}}{|z_1| \cdot |z_2|}$$

gilt. Im Unterschied zur Modellierung von Partikelbahnen kommt beim Wellenmodell also noch ein Interferenzterm hinzu. Die Größe des Interferenzterms hängt nach Abbildung 3 vom Differenzwinkel $\alpha - \beta = \Theta$ und damit von den Phasen der beiden Wegkomponenten ab. Im Aufbau von Abbildung 2 kann diese Phasendifferenz durch den Halbspiegel m_0 manipuliert werden.

5.3 Quantenmechanisches Entscheidungsmodell

Nach den Gesetzen der Wahrscheinlichkeitsrechnung gilt für den Zielzustand A bei gegebenem Startzustand S

$$p(A|S) = p(G|S) \cdot p(A|G) + p(U|S) \cdot p(A|U)$$

Gemäß Abbildung 2 handelt es sich dabei um die Summe der Pfadgewichte entlang der beiden Wege $S \rightarrow G \rightarrow A$ und $S \rightarrow U \rightarrow A$. Für die formale Übertragung in das physikalische Pfadmo-

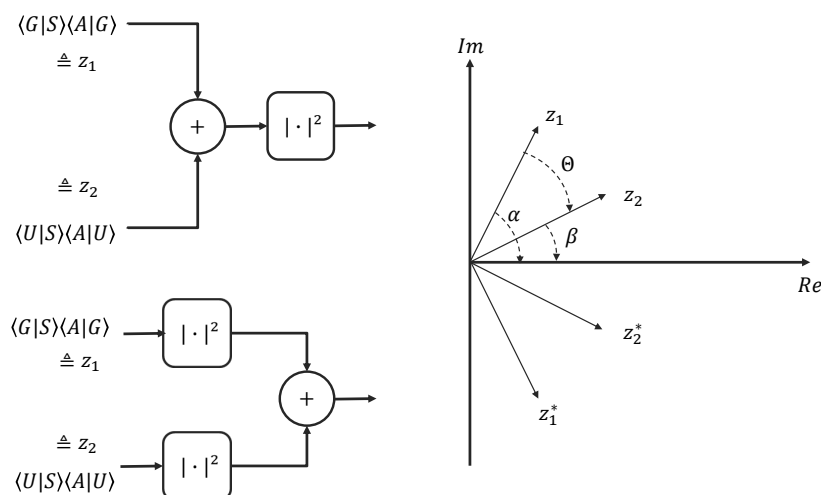


Abbildung 3 - Pfad- und Wellenmodell. Links unten: Im Pfadmodell werden die einzelnen Pfadwahrscheinlichkeiten addiert. Links oben: Wellenausbreitung mit Interferenzterm, hier werden die (komplexen) Amplituden der Wahrscheinlichkeitswelle addiert und anschließend quadriert. Rechts: Die Größe des Interferenzterms hängt von den komplexen Amplituden der zwei möglichen Pfade ab. Nur für $\Theta = \pi/2$ verschwindet die Interferenz.

dell verwenden wir für die Wahrscheinlichkeitsamplituden die Notation $\sqrt{p(A|S)} = \langle A|S \rangle$ und erlauben Belegungen mit komplexen Zahlen

$$|\langle A|S \rangle|^2 = |\langle G|S \rangle \langle A|G \rangle|^2 + |\langle U|S \rangle \langle A|U \rangle|^2.$$

Entsprechend gilt für das Wellenmodell

$$|\langle A|S \rangle|^2 = |\langle G|S \rangle \langle A|G \rangle|^2 + |\langle U|S \rangle \langle A|U \rangle|^2 + 2 \cdot |\langle G|S \rangle \langle A|G \rangle| \cdot |\langle U|S \rangle \langle A|U \rangle| \cdot \cos(\Theta).$$

Um die Forderung nach Veridikalität zu erfüllen [1], sollten die verwendeten Modelle die Wahrscheinlichkeiten der Ergebnisse von Experimenten bestätigen. Im Fall der klassischen Markov-Modelle, muss das Gesetz der totalen Wahrscheinlichkeit eingehalten werden, d.h. mit

$$p(G) = \sum_{s \in S} p(G|S)p(S) \text{ und } p(U) = \sum_{s \in S} p(U|S)p(S)$$

muss

$$p_T(A) = p(A|G)p(G) + p(A|U)p(U)$$

bestätigt werden. Wenn das Gesetz nicht erfüllt wird, so kann eine Anpassung über den Interferenzterm – insbesondere mit $\cos(\Theta)$ – erfolgen. Der Gewinn an Veridikalität ist dann allerdings mit Aufgabe der klassischen Sichtweise und der Verwendung eines komplexeren Modells zu bezahlen.

6 Diskussion und Ausblick

Die Lösung des Mouse-Maze-Problems ist ein Beispiel für produktives Denken oder Werkzeugintelligenz. Der elektromechanischen Maus „Theseus“ gelingt die Lösung — d.h. den Ausgang aus dem Labyrinth zu finden — durch die Benutzung eines physischen Werkzeugs: dem „elektromagnetischen“ Ariadnefaden. Bei der hier vorgeschlagenen quantenmechanischen Sichtweise auf die konsummatorische Endhandlung wird ein mentales Werkzeug benutzt: der virtuelle Doppelspalt. Formal wird dafür die zusätzliche Zufallsvariable C für die Kategorisierung und das Prinzip der Superposition benötigt. Der Nutzen eines solchen mentalen Werkzeugs besteht darin, dass der Agent nicht nur Eigenschaften des Zielobjekts messen kann sondern auch auf die subjektive Bedeutung von Zielobjekten schließen kann. Für ein finales System, ist dies eine besonders wichtige Fähigkeit. In diesem Beitrag wurden die Vorzüge quantenmechanischer Modelle nur in Bezug auf die Eigenschaft der Interferenz diskutiert, weitere Aspekte wie z.B. die Berücksichtigung von Reihenfolge-Effekten sollen zukünftig untersucht werden. Da das quantenmechanische Modell allein auf der Wahrscheinlichkeitstheorie beruht, erscheint auch hier die Kombination mit der Nutzentheorie sinnvoll.

Literatur

- [1] BISCHOF, N.: *Psychologie, ein Grundkurs für Anspruchsvolle*. Verlag Kohlhammer, 2009. 2. Auflage.
- [2] RÖMER, R. UND WIRSCHING, G.: *Ein Beitrag zu den natur- und geisteswissenschaftlichen Grundlagen kognitiver Systeme*. In: TODO und ET AL. (Herausgeber): *24. Konferenz Elektronische Sprachsignalverarbeitung, Bielefeld 2013*, Seiten 93 – 102. TUDpress, 2013.
- [3] RÖMER, R. UND WOLFF, M.: *Konzeption eines kognitiven Systems für den experimentellen Einsatz in Forschung und Lehre*. In: TODO und ET AL. (Herausgeber): *26. Konferenz Elektronische Sprachsignalverarbeitung, Eichstätt 2015*, Seiten 212 – 223. TUDpress, 2015.
- [4] WOLFF, M.; HUBER, M.; WIRSCHING G. RÖMER R. und I. SCHMITT: *Towards A Quantum Mechanical Model of the Inner Stage of Cognitive Agents*. In: *9th IEEE International Conference on Cognitive Infocommunications, Budapest 2018*, 2018.
- [5] WOLFF, M.; RÖMER, R. und G. WIRSCHING: *Towards coping and imagination for cognitive agents*. In: *6th IEEE International Conference on Cognitive Infocommunications, Győr 2015*, 2015.
- [6] BISCHOF, N.: *Struktur und Bedeutung, Einführung in die Systemtheorie*. Hogrefe, vorm. Verlag Hans Huber, 2016. 3. Auflage.
- [7] BUSEMEYER, J.R. und P.D. BRUZA: *Quantum Models of Cognition and Decision*. Cambridge University Press, 2012.
- [8] ZEILINGER, A.: *Einsteins Schleier*. Wilhelm Goldmann Verlag, München, 2005.
- [9] POLANI, D.: *Information: currency of life?* HFSP Journal, Vol. 3, No.5: pp. 307 – 316, 2009.
- [10] BLUTNER, R. und P. BEIM GRABEN: *Quantum cognition and bounded rationality*. Synthese, Vol. 193: pp. 3239 – 3291, 2016.