



Marco Ragni · Ute Schmid

# Kognitive Künstliche Intelligenz

MOREMEDIA



Springer

# Kognitive Künstliche Intelligenz

Marco Ragni • Ute Schmid

# Kognitive Künstliche Intelligenz



Springer

Marco Ragni  
Professur Prädiktive Verhaltensanalyse  
Technische Universität Chemnitz  
Chemnitz, Deutschland

Ute Schmid  
Kognitive Systeme  
Otto-Friedrich-Universität Bamberg  
Bamberg, Deutschland

ISBN 978-3-662-69497-8

ISBN 978-3-662-69498-5 (eBook)

<https://doi.org/10.1007/978-3-662-69498-5>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2025

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jede Person benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des/der jeweiligen Zeicheninhaber\*in sind zu beachten.

Der Verlag, die Autor\*innen und die Herausgeber\*innen gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autor\*innen oder die Herausgeber\*innen übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Einbandabbildung: © Oksana Tryndiak / Generated with AI / Stock.adobe.com

Planung/Lektorat: Marion Krämer

Springer ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Wenn Sie dieses Produkt entsorgen, geben Sie das Papier bitte zum Recycling.

# Vorwort

---

Das vorliegende Buch basiert auf dem 1996 erschienenen Werk *Kognitive Modellierung: Eine Einführung in logische und algorithmische Grundlagen* von Ute Schmid und Martin Kindsmüller. Ziel des Buches war es, Studierenden und Forschenden aus dem Bereich Kognitionspsychologie eine Einführung in zentrale Methoden der Künstlichen Intelligenz zu geben. Im Gegensatz zu Standardlehrbüchern zur Künstlichen Intelligenz werden die Themen so eingeführt, dass keine Vorkenntnisse im Bereich Informatik notwendig sind. Alle Themen werden zudem mit expliziten Bezügen zur kognitionswissenschaftlichen Forschung motiviert, und es wird besonderes Augenmerk auf viele klassische Grundlagenartikel gelegt, ohne die die aktuelle Entwicklung kaum verständlich wäre. Das Buch erschien im letzten der sogenannten KI-Winter – also in einer Zeit, in der kaum Interesse an Künstlicher Intelligenz (KI) außerhalb der engeren Forschungsgemeinschaft bestand. Dies hat sich mit dem Aufkommen neuer Ansätze, insbesondere *Deep Learning* und *Generative KI*, schlagartig geändert. Das Thema KI erfährt eine nie dagewesene Aufmerksamkeit und wird in nahezu allen wissenschaftlichen Disziplinen diskutiert, so auch in der Psychologie. Zunehmend werden Methoden der KI in vielen Bereichen angewendet – in der Psychologie als Ergänzung zu statistischen Methoden, zur Modellierung kognitiver Prozesse oder als Untersuchungsgegenstand.

Entsprechend war es an der Zeit, das oben genannte Buch in gründlich überarbeiteter und deutlich aktualisierter Form neu aufzulegen. Anders als 1996 ist die Bezeichnung „Künstliche Intelligenz“ nun wieder salonfähig und muss nicht hinter der Bezeichnung „Kognitive Modellierung“ versteckt werden. Die Initiative für das vorliegende Buch *Kognitive Künstliche Intelligenz* kam von Marco Ragni, der die Überarbeitung des ursprünglichen Werkes, bei dem Ute Schmid Erstautorin war, maßgeblich umgesetzt hat. Beide Autoren hoffen, dass dieses Buch Studierenden, Forschenden sowie Praktikerinnen und Praktikern einen hilfreichen Einstieg in Themen und Methoden der Künstlichen Intelligenz bietet.

Wir bedanken uns bei Christopher von Bülow für die kritische Durchsicht und finale Editierung des Manuskripts. Dominik Bär danken wir herzlich für die vielfältige Unterstützung, einschließlich der Bildbearbeitung. Ebenso sind wir Sara Todorovikj und Milena Stella Jans für ihre Unterstützung bei der Bildbearbeitung verbunden. Unser Dank geht besonders auch an unsere Lektorin beim Springer-Verlag, Marion Krämer, für ihre Geduld und Unterstützung.

**Marco Ragni**

**Ute Schmid**

Juli 2025

Zusätzliche Online-Materialien inkl. Dozentenfoliensatz zu diesem Buch finden Sie auf: ► <https://www.lehrbuch-psychologie.springernature.com>

# Inhaltsverzeichnis

---

1	<b>Einführung</b> .....	1
2	<b>Was ist Kognitive Künstliche Intelligenz?</b> .....	5
2.1	Der Informationsverarbeitungsansatz .....	8
2.2	Philosophische Grundlagen der Kognitionsforschung .....	10
2.3	Die Methode der kognitiven Modellierung .....	13
2.4	Zur Vertiefung .....	15
<b>I</b>	<b>Wissensrepräsentation und Logik</b>	
3	<b>Grundlagen der Wissensrepräsentation</b> .....	19
3.1	Begriffliches Wissen .....	22
3.2	Strukturiertes Wissen: semantische Netze .....	25
3.3	Zur Vertiefung .....	32
4	<b>Aussagenlogik</b> .....	33
4.1	Grundlagen .....	35
4.2	Syntax der Aussagenlogik .....	35
4.3	Semantik der Aussagenlogik .....	37
4.4	Äquivalenz aussagenlogischer Formeln .....	39
4.5	Schlussregeln .....	40
4.6	Zur Vertiefung .....	43
5	<b>Prädikatenlogik</b> .....	45
5.1	Syntax der Prädikatenlogik .....	47
5.2	Semantik von prädikatenlogischen Formeln .....	49
5.3	Formalisierung semantischer Netze .....	52
5.4	Zur Vertiefung .....	56
6	<b>Schlussfolgern und Beweisen</b> .....	59
6.1	Wozu beweisen? .....	60
6.2	Standardisierte logische Darstellungen .....	62
6.3	Logische Transformationsverfahren .....	66
6.4	Resolution – ein Schlussfolgerungsmechanismus .....	67
6.5	Schlussfolgerungen über Faktenwissen .....	71
6.6	Zur Vertiefung .....	74
7	<b>Logische Programmierung</b> .....	75
7.1	Was ist Programmierung? .....	76
7.2	Die Syntax von Prolog .....	78
7.3	Prolog und Prädikatenlogik .....	84
7.4	Zur Vertiefung .....	88

## II Kognition und Modellierung

8	<b>Algorithmen und formale Sprachen</b>	91
8.1	Problemlöseprozesse als Algorithmen	92
8.2	Algorithmen als Turing-Maschinen	97
8.3	Formale Sprachen	102
8.4	Zur Vertiefung	107
9	<b>Problemrepräsentation</b>	109
9.1	Listen, Bäume, Graphen	110
9.2	Probleme als Zustandsräume	116
9.3	Zur Vertiefung	123
10	<b>Allgemeine Suchstrategien und Komplexität</b>	125
10.1	Tiefensuche	127
10.2	Breitensuche	129
10.3	Aufwand, Komplexität und Berechenbarkeit	133
10.4	Zur Vertiefung	141
11	<b>Heuristiken</b>	143
11.1	Heuristische Suchstrategien	144
11.2	Problemlösen mit Constraints	154
11.3	Zur Vertiefung	157
12	<b>Kognitive Architekturen</b>	159
12.1	Kognitive Modelle in der KI-Forschung	160
12.2	Grenzen und Möglichkeiten der kognitiven Modellierung	163
12.3	Grundlagen von Produktionssystemen	167
12.4	Die Produktionssysteme GPS und Soar	174
12.5	Die kognitive Architektur ACT-R	180
12.6	Zur Vertiefung	186
13	<b>Lernen von Regeln</b>	189
13.1	Konzepterwerb	190
13.2	Entscheidungsbaum und Klassifizierungsfunktion	192
13.3	Zur Vertiefung	199
14	<b>Lernen von implizitem Wissen</b>	201
14.1	Aufbau und Arbeitsweise eines künstlichen Neurons	202
14.2	Aufbau und Arbeitsweise eines neuronalen Netzes	205
14.3	Zur Vertiefung	211

## III Ausgewählte Anwendungen

15	<b>Lernen und Expertise</b>	215
15.1	Lernen und Wissenserwerb im Überblick	216
15.2	Fertigkeitserwerb	218
15.3	Struktur und Erfassung von Expertenwissen	222

15.4	<b>Architektur von Expertensystemen</b> .....	226
15.5	<b>Zur Vertiefung</b> .....	230
16	<b>Intelligente Tutorsysteme</b> .....	233
16.1	<b>Design Intelligenter Tutorsysteme</b> .....	234
16.2	<b>Künstliche Intelligenz in der Bildung</b> .....	241
16.3	<b>Zur Vertiefung</b> .....	242
17	<b>Sprachverarbeitung: Syntaxanalyse</b> .....	245
17.1	<b>Aspekte der Sprachverarbeitung</b> .....	247
17.2	<b>Syntaxanalyse: Grammatik und Parser</b> .....	249
17.3	<b>Zur Vertiefung</b> .....	260
18	<b>Sprachverarbeitung: semantische Analyse</b> .....	261
18.1	<b>Bedeutung als Sinn und Referenz</b> .....	262
18.2	<b>Lexikalische und strukturelle Semantik</b> .....	263
18.3	<b>Zur Vertiefung</b> .....	267
19	<b>Mentale Modelle beim Textverstehen</b> .....	269
19.1	<b>Repräsentation von Wortbedeutung</b> .....	270
19.2	<b>Semantische Analyse sprachlicher Ausdrücke</b> .....	275
19.3	<b>Zur Vertiefung</b> .....	279
20	<b>Sprachverstehen: Von ELIZA zu Transformermodellen</b> .....	281
20.1	<b>ELIZA – der erste Chatbot</b> .....	282
20.2	<b>Watson</b> .....	285
20.3	<b>Transformermodelle</b> .....	286
20.4	<b>Zur Vertiefung</b> .....	289
21	<b>Ein Ausblick – Wie geht es weiter?</b> .....	291
	<b>Serviceteil</b> .....	295
	Glossar .....	296
	Literatur .....	328





# Einführung

Kognitive Künstliche Intelligenz (KKI) ist eine junge interdisziplinäre Wissenschaft, die sich mit natürlichen wie künstlichen, konkreten wie abstrakten kognitiven Prozessen auseinandersetzt und prädiktive<sup>1</sup> sowie erklärbare algorithmische Modelle<sup>2</sup> für diese entwickelt.

Sie stellt die formale Nachbardisziplin zur Kognitionswissenschaft dar, die ihrerseits von Kognitionspsychologie, Neurowissenschaft, Linguistik und analytischer Philosophie inspiriert ist. Kognition beim Menschen bezeichnet dabei alle mentalen, d. h. geistigen, Prozesse, die für die Verarbeitung von Information relevant sind. Dies umfasst deren Erwerb (durch Wahrnehmungsprozesse und Aufmerksamkeit), ihre spezifische mentale Repräsentation (z. B. im menschlichen Arbeitsgedächtnis), das Schlussfolgern (also die Verarbeitung der repräsentierten Information), das Anpassen vorhandenen Wissens (durch Lernen), die Verknüpfung von Informationen (Analogie), das Gewinnen von Einsichten bei neuen Aufgabenstellungen (Problemlösen) und die Kommunikation mit der Umwelt (z. B. Sprache; Kluwe 2000). Die Kognitionswissenschaft beschreibt menschliche kognitive Prozesse durch Modelle über menschliche Informationsverarbeitung (vgl. schon Strube 1993). Somit bildet die Kognitionswissenschaft einen wichtigen Baustein der KKI. Zugleich geht die KKI über die Kognitionswissenschaft hinaus, indem sie die kognitiven Konzepte nicht nur als Grundlage von KI betrachtet, sondern beobachtete kognitive Prozesse formalisiert und systematisch weiterentwickelt. Wesentliche Forschungsthemen beider Nachbardisziplinen sind alltägliche Ausprägungen von Intelligenz wie Wahrnehmung, Denken, Problemlösen, Sprache, Wissensorganisation und Lernen.

Kognitive Künstliche Intelligenz setzt eine interdisziplinäre Zusammenarbeit voraus, bei der gleichzeitig die einzelnen Disziplinen ihre spezifischen Schwerpunkte und Forschungsmethoden beibehalten. Die Verzahnung der von den verschiedenen Disziplinen verwendeten Methoden bringt ein mächtigeres Instrumentarium zur Erforschung kognitiver Prozesse hervor, als es jede Einzeldisziplin liefern könnte. Auch die Theoriebildung in den genannten Disziplinen ist häufig kognitiv orientiert. So sind viele kognitive Theorien ursprünglich von klassischen Ansätzen der KI beeinflusst, und die Ergebnisse kognitionspsychologischer Arbeiten beeinflussten und beeinflussen die Entwicklung der Künstliche-Intelligenz-Forschung.

Die interdisziplinäre Verankerung der KKI in Forschungsarbeiten macht sich zwangsläufig in der Lehre an der Universität bemerkbar. Beispielsweise werden bereits in einflussreichen Lehrbüchern zur kognitiven Psychologie (z. B. Anderson 2007b) Ansätze vorgestellt, die ihre Wurzeln in der KI haben oder von Arbeiten aus der KI beeinflusst sind. Insbesondere gilt dies für die Themen Wissensrepräsentation und Problemlösen. Der informatische Hintergrund dieser Ansätze wird zwar häufig erwähnt, die formalen Methoden, auf denen sie basieren, werden jedoch meist nicht eingeführt.

Ziel dieses Buches ist es, Prinzipien, Intuitionen, Ansätze und formale Methoden der KI zu vermitteln, die sich zur Modellierung kognitiver Prozesse als erfolgreich herausgestellt haben. Dabei haben wir uns bemüht, den Stoff so aufzubereiten, dass er auch von Lesern ohne Vorkenntnisse in Mathematik und Informatik nachvollzogen werden kann.

1 Prädiktive Modelle leiten aus vergangenen Beobachtungen Muster ab, um zukünftige Entwicklungen oder Verhalten vorherzusagen.

2 Erklärbare algorithmische Modelle (auch manchmal erklärbare KI) zielen darauf ab, die Funktionsweise komplexer KI-Modelle – insbesondere solcher, die nicht von Natur aus leicht verständlich sind (wie tiefes Lernen) – verständlich zu machen. Dadurch kann man als Mensch besser nachvollziehen, warum ein Modell bestimmte Entscheidungen trifft.

Es werden zunächst die wissenschaftstheoretischen und philosophischen Grundlagen der KKI dargestellt (► Kap. 2).

Danach ist das Buch in drei Teile gegliedert: Im ersten Teil des Buches werden formale Grundlagen für psychologische und kognitive Modellvorstellungen zur Repräsentation von deklarativem Wissen vermittelt. Hier beziehen wir uns vor allem auf den Ansatz der semantischen Netze, gehen aber auch kurz auf den schematheoretischen Ansatz ein. Beide Ansätze haben ihr Fundament in der formalen Logik. Als Grundlage für das Verständnis logischer Formalisierung führen wir zunächst wesentliche Konzepte der Mengenlehre ein (► Kap. 3). Wir erklären, wie hierarchische semantische Netze durch mengentheoretische Konzepte beschrieben werden können. In ► Kap. 4 und 5 stellen wir die Syntax der Aussagen- und der Prädikatenlogik vor und zeigen die Anwendbarkeit logischer Schlussregeln. Wir veranschaulichen dann, wie durch Überführung natürlichsprachiger Sätze in die Syntax der Aussagenlogik semantische Netze konstruiert werden können. Die wichtigste Methode zum Schlussfolgern aus gegebenem Wissen ist der Theorembe-  
weis. Diesen Ansatz stellen wir in ► Kap. 6 dar. Die Prädikatenlogik ist die wesentliche Grundlage für die Programmiersprache Prolog (► Kap. 7). Wir führen Prolog beispielhaft anhand der Implementation eines semantischen Netzes ein.

Der zweite Teil des Buches dient der Vermittlung der Grundlagen für die kognitive Modellierung von Problemlösefertigkeiten. Hier stellen wir zunächst die Grundbegriffe der theoretischen Informatik dar; insbesondere führen wir Algorithmen und formale Sprachen ein (► Kap. 8). Viele Ansätze zur Modellierung menschlicher Denk- und Problemlöseprozesse basieren auf diesen Grundlagen. Im nächsten Kapitel (► Kap. 9) stellen wir dar, wie Probleme repräsentiert werden können, und zeigen, dass Problemlösen als Suche in einem „Problemraum“ beschrieben werden kann. Wir stellen zwei grundlegende Suchverfahren vor (► Kap. 10), die dann in ► Kap. 11 zu heuristischen Suchverfahren erweitert werden. Heuristische Suchverfahren sind ein zentraler Bestandteil von Produktionssystemen, die wir in ► Kap. 12 einführen. Zum Abschluss des zweiten Teils werden wir uns mit dem Lernen von Regeln (► Kap. 13) beschäftigen sowie eine kurze Einführung in den Bereich der neuronalen Netze geben (► Kap. 14).

Ziel des dritten Teils ist es, die Anwendung der vermittelten Grundlagen in verschiedenen Forschungsbereichen aufzuzeigen. Exemplarisch stellen wir die Bereiche Lernen und Expertise (► Kap. 15 und 16), Sprachverarbeitung (► Kap. 17 und 18) sowie Textverstehen (► Kap. 19 und 20) dar. Das Buch schließt mit einem kurzen Ausblick und einer Einschätzung des Potentials und der Grenzen der dargestellten Ansätze zur Modellierung kognitiver Prozesse (► Kap. 21) und greift damit die in ► Kap. 2 dargestellten Themen wieder auf.

Wir legen vor allem Wert auf eine umfassende und verständliche Einführung der formalen Grundlagen. Aus diesem Grund stellen wir nicht nur die aktuellen Ansätze vor, sondern beziehen uns wann immer möglich auf die „Klassiker“, welche oftmals viele Entwicklungen vorweggenommen haben. Am Ende jedes Kapitels geben wir Hinweise auf weiterführende Arbeiten. Deutsche Übersetzungen von Fachtermini verwenden wir immer, wenn die entsprechenden Begriffe in der deutschsprachigen Literatur gebräuchlich sind. Wo dies nicht der Fall ist, geben wir den englischen Begriffen den Vorzug, auch wenn dadurch unschöne Kombinationen aus englischen und deutschen Worten entstehen. Wir fokussieren auf die Darstellung symbolischer Ansätze, gehen aber auch auf die Prinzipien der Modellierung mit neuronalen Netzen ein.

Wir gehen in mehreren Kapiteln auf die Umsetzung der dargestellten Formalismen in Computerprogrammen ein. Da es uns nicht als sinnvoll erscheint, die Einführung formaler

Konzepte durch die Erläuterung technischer Begriffe zu unterbrechen, führen wir die Definitionen informatischer Konzepte in einem Glossar auf.

Unser Anliegen ist, dass nach Lektüre dieses Buches Spezialliteratur zur kognitiven Modellierung sowie Lehrbücher zur Künstlichen Intelligenz ohne große Verständnisprobleme gelesen werden können. Das Buch ist so aufgebaut, dass die in Teil I und II dargestellten Inhalte in einer einsemestrigen Lehrveranstaltung vermittelt werden können. Die in Teil III dargestellten Themen sollen Anregungen für die Auseinandersetzung mit speziellen Forschungsgebieten der Kognitionswissenschaft geben. Hier empfiehlt es sich unserer Meinung nach, spezielle Veranstaltungen zu konzipieren, die über die von uns dargestellten Inhalte hinausgehen.



# Was ist Kognitive Künstliche Intelligenz?

## Inhaltsverzeichnis

- 2.1 Der Informationsverarbeitungsansatz – 8
- 2.2 Philosophische Grundlagen der Kognitionsforschung – 10
- 2.3 Die Methode der kognitiven Modellierung – 13
- 2.4 Zur Vertiefung – 15

„Kognitive Künstliche Intelligenz“ ist ein Sammelbegriff, unter dem die Forschung zu kognitiven Strukturen, Prozessen und wissensbasierten technischen Systemen zusammengefasst wird. Klassische Forschungsthemen sind Wahrnehmung, Denken, Problemlösen, Sprache, Wissensorganisation und Lernen. Disziplinen, in denen Fragestellungen aus diesen Bereichen bearbeitet werden, sind insbesondere Künstliche Intelligenz, kognitive Psychologie, Neurowissenschaften, Linguistik und Philosophie.

Zentrales Forschungsanliegen der Künstlichen Intelligenz ist der Entwurf von Algorithmen, die komplexe Probleme aus den oben genannten Bereichen bewältigen können. In der kognitiven Psychologie werden Modelle menschlicher Kognition aufgestellt und empirisch überprüft. Die Neurowissenschaften beschäftigen sich mit Aufbau und Funktionsweise von Nervensystemen, insbesondere des Gehirns, und erforschen diese empirisch. Die Linguistik wiederum definiert normative und deskriptive Modelle über die Regularitäten sprachlicher Strukturen. Und in der Philosophie werden Grundlagen und Grenzen des Denkens und Verhaltens analysiert.

Kognitive Strukturen und Prozesse werden in jeder Disziplin mit deren eigenen Zielsetzungen und Methoden erforscht. Zum Teil haben sich bereits so starke interdisziplinäre Anknüpfungspunkte zwischen je zwei Disziplinen ausgebildet, dass sich Forschungsgebiete wie Neuropsychologie, Neuroinformatik, Computerlinguistik und Psycholinguistik herausgebildet haben.

In diesem Buch wird insbesondere die Interaktion von kognitiver Psychologie und Künstlicher Intelligenz behandelt. Der Schnittbereich dieser beiden Disziplinen ist die Modellierung kognitiver Prozesse. Sowohl kognitive Psychologie als auch Künstliche Intelligenz sind von ihrem Ursprung her stark interdisziplinär orientiert. Die Wende der Psychologie in den 50er-Jahren des vergangenen Jahrhunderts vom Behaviorismus hin zum Paradigma der Kognitionsforschung wurde unter anderem durch das Aufkommen der Kommunikations- und Informationstheorie (Shannon 1948) und der Kybernetik (Wiener et al. 2019) sowie von den Arbeiten des Linguisten Chomsky eingeleitet. Kommunikationstheorie, Informationstheorie und Kybernetik bildeten gleichzeitig wesentliche Grundlagen für die in den 60er-Jahren des 20. Jahrhunderts entstehende Informatik.

Als sich entwickelnde Teildisziplin der Informatik entstand die Künstliche Intelligenz. So wurde der Begriff „artificial intelligence“ (deutsch: Künstliche Intelligenz, kurz KI) erstmalig 1956 auf einer Computer-Konferenz<sup>1</sup> verwendet. Zu den „Gründervätern“ der KI gehörten neben Alan Turing, John McCarthy und Marvin Minsky auch Forscher wie Allen Newell und Herbert A. Simon, die die kognitive Psychologie entscheidend mitgeprägt haben.

Das gemeinsame Interesse der KI und der kognitiven Psychologie an der Modellierung menschlicher Kognition führte jedoch nicht zum Aufgehen beider Disziplinen in der Kognitionsforschung (Schmalhofer und Wetter 1988). Die kognitive Psychologie hat primär das Ziel, menschliche Informationsverarbeitungsprozesse zu erforschen. Eine wesentliche Methode, sowohl zur Bildung als auch zur Prüfung von Hypothesen, ist dabei die empirische Untersuchung der kognitiven Leistungen des Menschen. Die Künstliche Intelligenz konzentriert sich darauf, effiziente Systeme zu entwickeln, die komplexe Aufgaben bewältigen können, die ursprünglich nur durch menschliche Intelligenz lösbar waren. Ein Beispiel dafür sind Expertensysteme, die bereits in den 1970er-Jahren entwickelt wurden, um menschliche Experten durch die Abbildung ihres Wissens und

<sup>1</sup> „The Dartmouth Summer Research Project on Artificial Intelligence“, organisiert von McCarthy. Unter den Vortragenden waren neben McCarthy unter anderem Minsky und Newell.

ihrer Schlussfolgerungsmechanismen zu simulieren. Geht man dabei von der Annahme aus, dass menschliche Informationsverarbeitung effizient ist, so ist es für die Künstliche Intelligenz von Interesse, beim Entwurf ihrer Systeme Erkenntnisse über menschliche Informationsverarbeitungsprozesse zu berücksichtigen.

Andererseits unterliegt die psychologische Theoriebildung – wie alle wissenschaftlichen Theorien – den Kriterien der Konsistenz und Prüfbarkeit (Schneewind 1977). Konsistente Theoriebildung ist jedoch nur durch eine hinreichend präzise sprachliche Darstellung der Konzepte möglich. Natürlichsprachige Darstellungen von Theorien sind oftmals unpräzise (Westmeyer 1977): Einerseits werden in verschiedenen Theorien über denselben Gegenstand unterschiedliche Begriffe verwendet, andererseits können in verschiedenen Theorien unterschiedliche Ideen mit denselben Begriffen ausgedrückt werden. Dies macht die Theorien schwer miteinander vergleichbar. Oft ist es sogar schwierig, für eine einzige Theorie zu prüfen, ob die ihr zugrunde liegenden Annahmen widerspruchsfrei sind und ob alle Annahmen, die notwendig sind, um den behandelten Gegenstand zu beschreiben, auch tatsächlich explizit formuliert wurden. So lässt sich menschliches Vergessen in der sogenannten artikulatorischen Schleife (also dem Gedächtnissystem, welches uns hilft, verbale Information wie Telefonnummern uns zu merken) auf 144 verschiedene Arten implementieren (Lewandowsky und Farrell 2011).

Mit einer formalen Darstellung von Theorien können diese Probleme ausgeräumt werden (Konerding 1992). Unter Formalisierung versteht man, dass Theorien in einer formalen Sprache mit streng festgelegter Syntax und Semantik dargestellt werden. Alle Aussagen einer Theorie könnten zum Beispiel in der Syntax der Prädikatenlogik (► Kap. 5) formuliert werden. Zu jedem eingeführten Symbol in dieser Sprache muss zudem genau festgelegt werden, was es bezeichnen soll. Gleiche Symbole bezeichnen dann immer exakt gleiche Konzepte.

In den 70er-Jahren des 20. Jahrhunderts wurde im Rahmen der mathematischen Psychologie versucht, psychologische Theorien zu formalisieren (Krantz et al. 1974). Mathematische Modelle sind jedoch in den meisten Fällen statisch. Sie sind sehr gut dazu geeignet, Beziehungen zwischen vorliegenden Stimuli und resultierendem Verhalten formal abzubilden.

Eine die Mathematik ergänzende formale Sprache liefert das Berechnungsmodell der Informatik (► Kap. 8), bei dem Verarbeitungsprozesse im Vordergrund stehen. Die Betonung der Prozesskomponente in den Modellen der Informatik macht diese besonders gut geeignet, Eigenschaften der Informationsverarbeitung formal zu beschreiben. Die kognitionswissenschaftliche Theoriebildung sollte daher beide genannten Aspekte berücksichtigen: Theorien über kognitive Strukturen und Prozesse sollten präzise und mithilfe einer formalen Sprache formuliert sein. Gleichzeitig sollte in den Theorien der erfahrungswissenschaftlichen Gegenstand, also die Aspekte menschlicher Kognition, über die in einer Theorie Aussagen getroffen werden, im Vordergrund stehen. Eine empirische Prüfung der Validität solcher Theorien ist also unverzichtbar.

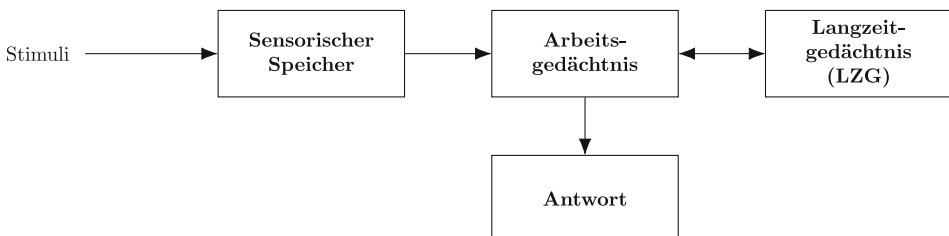
Im Folgenden werden wir die Bezeichnungen „Kognitionswissenschaft“ (engl. *cognitive science*) und „Kognitionsforschung“ synonym verwenden. Gehen wir speziell auf psychologische Beiträge zur Kognitionsforschung ein, so sprechen wir von „kognitiver Psychologie“ oder „Kognitionspsychologie“.

## 2.1 Der Informationsverarbeitungsansatz

Grundlage der meisten kognitionspsychologischen Theorien ist der Informationsverarbeitungsansatz, der den menschlichen Geist als ein System beschreibt, das Informationen ähnlich einem Computer in verschiedenen Stufen verarbeitet. Dieser Ansatz geht davon aus, dass Informationen aufgenommen, gespeichert, transformiert und abgerufen werden, was zentrale kognitive Prozesse wie Wahrnehmung, Gedächtnis und Problemlösen umfasst (etwa Lindsay und Norman 1977; Anderson 2000). Ein zentraler Ansatz, auf dem spätere Theorien aufbauen, ist das Multispeichermodell von Atkinson und Shiffrin (1968), das deutliche Parallelen zur Computerarchitektur<sup>2</sup> aufweist (■ Abb. 2.1).

Hier werden eingehende Informationen (**Stimuli**) in aufeinanderfolgenden Stufen verarbeitet, wobei diese Stufen sich wechselseitig beeinflussen können. Umweltinformationen werden zunächst für einen Sekundenbruchteil im sinnesspezifischen **sensorischen Speicher** gehalten. Ein Teil dieser Informationen gelangt ins **Arbeitsgedächtnis** (AG), wo physikalische Reize in symbolische Informationen umgewandelt werden. Die im Arbeitsgedächtnis repräsentierten Informationen stehen für aktuelle kognitive Prozesse zur Verfügung, jedoch ist dessen Kapazität begrenzt. Nur eine begrenzte Menge an Information kann gleichzeitig aktiv verarbeitet und genutzt werden, was die Effizienz der Informationsverarbeitung beeinflusst. Die restlichen Informationen werden vergessen, da das sensorische Gedächtnis nur eine sehr flüchtige Speicherung ermöglicht.

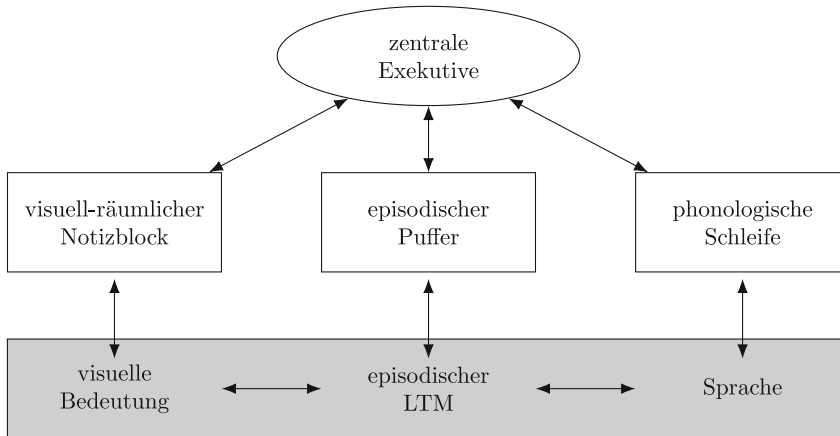
Das zentrale Papier hierzu („Magical Number Seven“; Miller 1956) hat eine weitere Schlüsselfigur, George Miller, geschrieben. In diesem Artikel wurde einer der ersten systematischen Versuche zur Modellierung der kognitiven Ressourcen des Arbeitsgedächtnisses beschrieben und zu einer Codierung im Arbeitsgedächtnis, dem Chunking, was sich z. B. in der Encodierung von Zahlen durch Muster abbilden lässt. Die Telefonnummer 2 4 6 1 3 5 können Sie sich leicht merken, denn es sind die ersten geraden Zahlen und dann die ersten ungeraden Zahlen. Die Rolle des Kurzzeitgedächtnisses wurde, wie oben erwähnt, durch das Arbeitsgedächtnis als ein Prozessmodell anstelle eines „statischen Speichers“ abgelöst (Baddeley 2000). Dieses Modell ist noch viel näher an der Computermetapher (vgl. ■ Abb. 2.2), in diesem Modell existieren eine Zentrale Exekutive (vergleichbar mit der CPU eines Computers), welche im Arbeitsgedächtnis die



■ **Abb. 2.1** Die stufenartige Verarbeitung von Informationen erfolgt nach dem Modell von Atkinson und Shiffrin (1968) über den sensorischen Speicher, das Kurzzeit- und das Langzeitgedächtnis. In diesem Modell wurde das Kurzzeitgedächtnis durch das heutige vorherrschende Arbeitsgedächtnis ersetzt (s. u.)

<sup>2</sup> Mit „Computerarchitektur“ ist hier speziell die Architektur des sogenannten Von-Neumann-Rechners gemeint. John von Neumann schlug 1946 ein Konzept zur Gestaltung eines universellen Rechners vor, an dem sich die meisten modernen Computer orientieren.





■ **Abb. 2.2** Die aktuelle Version des Mehrkomponenten-Arbeitspeichermodells nach Baddeley (2000) umfasst die Zentrale Exekutive, den visuell-räumlichen Notizblock und die artikulatorische Schleife als Kernkomponenten des Arbeitsgedächtnisses. Baddeley ergänzte das Modell später um den episodischen Puffer. In grau sind Elemente des Langzeitgedächtnisses enthalten

Informationsverarbeitung steuert, und kleinere informationsspezifische Speicher, wie ein visuell-räumlicher Speicher und ein sprachbasierter Speicher, welcher zum Beispiel für das Merken von Zahlenfolgen eingesetzt wird, sowie ein episodischer Speicher, der visuelle und räumliche Informationen von Ereignissen inklusive von Erinnerungen aus dem Langzeitgedächtnis integrieren kann. Diese Aspekte könnten Sie an den RAM- oder Arbeitsspeicher eines Computers erinnern.

Zusätzlich zu Informationen, die über die Sinnesorgane ins Arbeitsgedächtnis gelangen, können dort auch aus dem **Langzeitgedächtnis (LZG)** aktivierte Informationen für die Verarbeitung bereitgehalten werden. Im Langzeitgedächtnis gespeicherte kognitive Strukturen wirken dabei auf die Kategorisierung und Wahrnehmung spezifischer sensorischer Reize ein. So sehen wir nicht alles, was das Auge wahrnimmt, sondern nur das, worauf wir unsere Aufmerksamkeit richten. Abhängig von aktuellen Zielen und weiteren Faktoren wird ein Teil der Informationen ins Langzeitgedächtnis übernommen (Lernen). Es wird angenommen, dass Wissen im Langzeitgedächtnis permanent gespeichert ist und Sie zum Beispiel an die Festplatte im Computer erinnern kann. „Vergessen“ ist dann auf ein Nichtauffinden von Information zurückzuführen, das nicht mit einem Löschen der Information gleichzusetzen ist. Zudem wird angenommen, dass Wissen im Langzeitgedächtnis nicht als unzusammenhängende Menge von Einzelinformationen, sondern strukturiert repräsentiert ist. Das Wissen im Langzeitgedächtnis wird häufig nach verschiedenen Aspekten unterschieden. Zum Beispiel unterscheidet Anderson (1996) zwischen der Repräsentation von Faktenwissen (deklaratives Wissen) und Regelwissen (prozedurales Wissen). In den folgenden Kapiteln werden verschiedene Konzeptionen für die Repräsentation von Wissen vorgestellt. Dabei liefern Annahmen und Erkenntnisse über die Charakteristika menschlicher Informationsverarbeitungsprozesse – wie zum Beispiel die beschränkte Kapazität des Arbeitsgedächtnisses – wichtige Randbedingungen für den Entwurf von Modellen.

## 2.2 Philosophische Grundlagen der Kognitionsforschung

In der Kognitionsforschung werden kognitive Prozesse als Informationsverarbeitungsprozesse beschrieben, was aber die Annahme eines grundlegenden erkenntnistheoretischen Paradigmas ist. Dieses Informationsverarbeitungsparadigma lässt sich wie folgt charakterisieren (siehe Stillings et al. 1998, zitiert nach Strube 1993, S. 304):

Annahme 1: Unterschiedliche kognitive Prozesse basieren auf gemeinsamen allgemeinen und grundlegenden Prinzipien.

Annahme 2: Kognition lässt sich unabhängig von ihrem materiellen Substrat, also der Neuronenaktivität, betrachten.

Annahme 3: Informationsverarbeitende Prozesse lassen sich als formale Prozesse beschreiben.

Annahme 4: Informationsverarbeitende Prozesse operieren auf Repräsentationen.

Eng damit verbunden ist die Annahme, dass sich dieses Informationsparadigma am besten durch einen Symbolverarbeitungsansatz modellieren lässt. Der oben bereits erwähnte KI-Forscher Allen Newell (1980) hat das in der „*physical symbol systems hypothesis*“ explizit formuliert (siehe auch Newell und Simon 1976).<sup>3</sup> Die Hypothese besagt, dass alle kognitiven Prozesse nichts anderes als **Transformationen von physikalischen Symbolstrukturen** sind. Für unsere Betrachtungen genügt die abgeschwächte Hypothese, dass sich kognitive Prozesse als Transformationen von Symbolstrukturen **beschreiben** lassen.

Symbolstrukturen<sup>4</sup> und Regeln für ihre Transformation liefern die Basis für eine einheitliche Beschreibung kognitiver Prozesse (► Annahme 1). Symbolstrukturen werden durch Verknüpfung elementarer Symbole aufgebaut, wobei die Verknüpfung syntaktischen Regeln gehorcht. Eine einfache Symbolstruktur wäre die folgende:

- Elementare Symbole:  $S = \{0, 1, \#\}$
- Symbolstrukturen  $E$  können beispielsweise durch folgende syntaktische Regeln  $R_i$  aus Symbolen aus  $S$  aufgebaut werden:
  - $R_1: E \rightarrow \#$ ,
  - $R_2: E \rightarrow 0E$ ,
  - $R_3: E \rightarrow 1E$ .

Dabei kann der Pfeil gelesen werden als „... wird ersetzt durch ...“. Beispielsweise wird durch die zweite Regel das Symbol  $E$  durch die beiden Symbole  $0E$  ersetzt. Damit können längere Zeichenketten erzeugt werden. Die Anwendung dieser Regeln ermöglicht es zum Beispiel, folgende Strukturen zu erzeugen:

$E \rightarrow \#$  (Anwendung von  $R_1$ )

$E \rightarrow 0E \rightarrow 0\#$  (Anwendung von  $R_2$  und  $R_1$ )

$E \rightarrow 1E \rightarrow 10E \rightarrow 100E \rightarrow 100\#$  (Anwendung von  $R_3$ ,  $R_2$ ,  $R_2$  und  $R_1$ )

Damit ist  $E$  beschränkt auf beliebige (auch leere) Folgen aus den Ziffern 0 und 1, die mit dem Symbol  $\#$  abgeschlossen sind. Das Symbol  $\#$  kann als Zeichen für die leere Zeichenfolge/Symbolstruktur interpretiert werden. **Formale Sprachen**, die durch Symbolfolgen

3 In ► Kap. 12 werden wir auf die durch die „*Physical symbol systems*“-Hypothese ausgelöste philosophische Debatte eingehen.

4 In der theoretischen Informatik wird dies auch als Wort (endliche Symbolfolge) bezeichnet.

und syntaktische Regeln definiert sind, spielen nicht nur in der Theorie, sondern auch in der Praxis der Informatik eine zentrale Rolle. Im Kontext von Computern werden solche Sprachen verwendet, um alle Arten von Informationen zu verarbeiten. Da Computer nur mit elektrischen Zuständen arbeiten können, die zwei Zustände („an“ und „aus“) darstellen, erfolgt die Codierung von Informationen im sogenannten binären Zahlensystem, das nur die Symbole 0 und 1 kennt. Das bedeutet, dass wir, wenn wir beliebige Symbolstrukturen mit dem Computer verarbeiten wollen, eine Übersetzung aller Symbolstrukturen in binäre Symbolstrukturen benötigen. Wir demonstrieren im Folgenden, wie wir Dezimalzahlen auf binäre Zahlen abbilden können.

Im Dezimalsystem stehen uns die Ziffern 0 bis 9 zur Verfügung. Zahlen werden aus Summen von Zehnerpotenzen gebildet. So ergibt sich die Dezimalzahl 86 als

$$86_{10} = 8 \cdot 10^1 + 6 \cdot 10^0.$$

Die Einerstelle wird durch Multiplikation der Ziffer mit  $1 = 10^0$  gebildet (für alle Zahlen  $x$  gilt  $x^0 = 1$ ). Die Zehnerstelle wird durch Multiplikation der Ziffer mit  $10^1 = 10$  gebildet, die Hunderterstelle durch Multiplikation der Ziffer mit  $10^2 = 100$  und so fort. Im binären Zahlensystem haben wir nur die Ziffern 0 und 1 zur Verfügung. Hier werden Zahlen aus Summen von Zweierpotenzen gebildet. So können wir die Dezimalzahl 86 durch

$$\begin{array}{cccccccc} 1 \cdot 2^6 & + & 0 \cdot 2^5 & + & 1 \cdot 2^4 & + & 0 \cdot 2^3 & + & 1 \cdot 2^2 & + & 1 \cdot 2^1 & + & 0 \cdot 2^0 & = \\ 64 & + & 0 & + & 16 & + & 0 & + & 4 & + & 2 & + & 0 & = 86_{10} \end{array}$$

als die Binärzahl 1010110 repräsentieren. Auch beliebige Zeichenfolgen, wie zum Beispiel Worte wie „Vogel“ oder „Hund“ lassen sich durch solche 0–1-Kombinationen ausdrücken. Im Computer werden solche Begriffe auf diese Art, im binären Zahlensystem, als Folge von ASCII-Zeichen (*American Standard Code for Information Interchange*) codiert: „V“ hat beispielsweise die ASCII-Code-Nummer 86; im binären Zahlensystem wird daraus 1010110. Im Folgenden verzichten wir meist auf diese schwer lesbare Darstellung und verwenden Worte oder Dezimalzahlen als Elementarsymbole.

Bisher haben wir gezeigt, wie aus elementaren Symbolen mithilfe von syntaktischen Regeln komplexere Symbolstrukturen aufgebaut werden können. Wir haben eine **Grammatik** angegeben, die den syntaktischen Aufbau von Symbolstrukturen festlegt. Verwenden wir als Elementarsymbole Worte, wie zum Beispiel

$$S = \{\text{Peter, Eva, lacht, tanzt}\},$$

und als Regeln

$$\begin{array}{ll} R_1: & \text{Satz} \rightarrow \text{Eigenname Verb}, \\ R_2: & \text{Eigenname} \rightarrow \text{Peter}, \\ R_3: & \text{Eigenname} \rightarrow \text{Eva}, \\ R_4: & \text{Verb} \rightarrow \text{lacht}, \\ R_5: & \text{Verb} \rightarrow \text{singt}, \end{array}$$

so können wir einfache Sätze wie „Peter lacht“ als Symbolstrukturen erzeugen. Um **Transformationen** von Symbolstrukturen zu beschreiben, benötigen wir weitere Regeln.

Diese Regeln müssen angeben, wie ein bestimmter Ausdruck  $E$  einer Symbolstruktur durch neue Ausdrücke ersetzt werden kann. Es könnte zum Beispiel eine Regel für die Addition definiert werden, die für Ausdrücke der Form  $x + y$  die Addition ausführt. Für  $E = 17 + 11$  würde die Regel also das Ergebnis 28 liefern. Die Symbolstruktur  $E$  wird durch Regelanwendung in eine neue Symbolstruktur, etwa  $E' = 28$ , transformiert. Regeln zur Symboltransformation liefern uns eine Möglichkeit, Informationsverarbeitungsprozesse wie etwa die Addition von Zahlen zu beschreiben. Umgangssprachlich formuliert könnte die Addition von zwei Zahlen, die aus je zwei Ziffern bestehen, durch folgende Regeln definiert werden:

- $R_1$ : Nimm von beiden Zahlen jeweils die hintere Ziffer und bilde ihre Summe und lösche die beiden Ziffern.
- $R_2$ : Notiere die letzte Ziffer der Summe als letzte Ziffer des Additionsergebnisses.
- $R_3$ : Besteht die Summe aus mehr als einer Ziffer, dann merke dir die noch nicht notierte Ziffer als Übertrag.
- $R_4$ : Nimm die vordere Ziffer der beiden Zahlen und, falls ein Übertrag existiert, den Übertrag, bilde ihre Summe und lösche die beiden Ziffern.
- $R_5$ : Notiere die letzte Ziffer der neuen Summe als vorletzte Ziffer des Additionsergebnisses.
- $R_6$ : Besteht die Summe aus mehr als einer Ziffer, dann notiere die noch nicht notierte Ziffer als vordere Ziffer des Additionsergebnisses.

In ► Kap. 8 und 9 werden wir eine spezielle Form von Transformationsregeln, sogenannte **Produktionsregeln**, einführen. Wie wir dort zeigen werden, eignen sich Produktionsregeln besonders gut zur Modellierung von Informationsverarbeitungsprozessen. Eine **Menge von Elementarsymbolen, Regeln zum Aufbau syntaktisch korrekter Symbolstrukturen** aus diesen Elementarsymbolen und **Regeln zu ihrer Transformation** definieren zusammen ein formales System zur Beschreibung von Informationsverarbeitungsprozessen, so wie es in Annahme 3 gefordert wird. Die Transformationsregeln operieren nicht auf konkreten Dingen der Welt, sondern auf Symbolstrukturen, die diese Dinge repräsentieren, so wie es in Annahme 4 gefordert wird.

Die zweite Annahme besagt, dass informationsverarbeitende Prozesse durch symbolische Codierungen beschrieben werden können, die unabhängig von der Materie (der „Hardware“) sind. In letzter Konsequenz heißt das auch, dass symbolische Codierungen Beschreibungen für die Informationsverarbeitungsprozesse verschiedener Systeme sind, egal ob diese Prozesse auf der „Maschine Gehirn“ oder dem Computer realisiert sind. Diese Auffassung wird nicht von allen Kognitionswissenschaftlern geteilt. Man kann durchaus moderatere Annahmen machen, wie zum Beispiel, dass Computersimulationen über Einzelaspekte menschlicher Denkprozesse Aufschluss geben können. Dabei ist es nicht notwendig, von einer Äquivalenz zwischen Computern und Menschen als symbolverarbeitende Systeme auszugehen, wohl aber von der Zulässigkeit, Informationsverarbeitungsprozesse unabhängig von den Prozessen des Gehirns zu beschreiben.

Die Annahme, dass geistige Prozesse unabhängig von ihrer materiellen Realisierung betrachtet werden können, liefert den Hintergrund für die **komputationale Theorie des Geistes**, eine wichtige Strömung der analytischen Philosophie. In dieser Theorie wird argumentiert, dass menschliche Denkprozesse auf **mental**en **Repräsentationen** operieren. Diese Repräsentationen können als syntaktisch strukturierte Symbole beschrieben wer-

den, die aus atomaren Elementen, den oben eingeführten Elementarsymbolen, aufgebaut sind. Die Bedeutung der Elementarsymbole wird dabei als gegeben vorausgesetzt. Ausgehend von diesen elementaren Bedeutungen lassen sich dann alle komplexeren Strukturen mithilfe von Regeln rein syntaktisch konstruieren. Fodor und Pylyshyn (1988) postulieren eine „Sprache des Geistes“ („**Mentalesisch**“), die nach diesen Regeln aufgebaut ist. Allgemein zeichnen sich kognitive Prozesse im Sinne der komputationalen Theorie des Geistes durch folgende Eigenschaften aus (Pylyshyn 1984; Strube 1993, S. 306):

1. **Produktivität:** Aus einer endlichen Menge atomarer Elemente kann eine unendliche Anzahl komplexer mentaler Ausdrücke generiert werden. Diese Annahme entspricht der Argumentation von Chomsky (2014) über das Produzieren und Verstehen natürlicher Sprache (► Kap. 17).
2. **Systematizität:** Wird eine bestimmte Satzstruktur verstanden, so verstehen wir alle Sätze dieser Struktur: Wenn man „John liebt Anna“ versteht, dann auch „Anna liebt John“ oder „Peter liebt Inge.“
3. **Kompositionalität:** Die Bedeutung komplexer Ausdrücke ergibt sich aus den Bedeutungen ihrer Bestandteile. Die Bedeutung von „Peter lacht“ ergibt sich also aus der Bedeutung von „Peter“ und der Bedeutung von „lacht“ zusammen mit der syntaktischen Relation, die zwischen den beiden atomaren Elementen besteht.

Diese kurze Einführung in den philosophischen Hintergrund der Kognitionsforschung sollte verdeutlichen, dass die Entscheidung für bestimmte Modellierungsmethoden immer von theoretischen Vorannahmen beeinflusst ist. Die Gültigkeit dieser Vorannahmen kann nie belegt werden; man kann sie entweder akzeptieren oder nicht. Die in den folgenden Kapiteln dargestellten logischen und algorithmischen Ansätze basieren auf den dargestellten Annahmen.

## 2.3 Die Methode der kognitiven Modellierung

Die Integration mehrerer Wissenschaften in der Kognitionsforschung führt, vor allem auf methodischer Ebene, zu einer Ergänzung und Erweiterung von Ansätzen, die in den Einzeldisziplinen zur Modellierung menschlicher Informationsverarbeitungsprozesse verwendet werden. Nach Strube (1993) lassen sich dabei drei Methodengruppen unterscheiden:

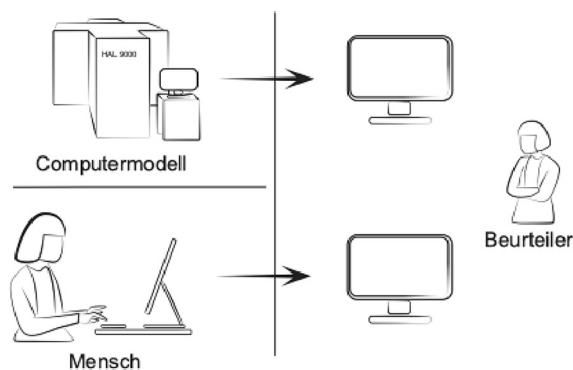
1. **Theoretische Analyse kognitiver Prozesse mit dem Ziel der Formalisierung:** Diese Methodologie stammt aus den Geistes- und Formalwissenschaften (Philosophie, Linguistik, Logik) und dient vor allem zur Bestimmung von Regularitäten und Beschränkungen mentaler Prozesse.
2. **Modellierung kognitiver Prozesse durch Computersimulation:** Diese Methodologie stammt aus der Informatik (Künstliche Intelligenz, Algorithmik, formale Sprachen) und dient vor allem dem Entwurf von Prozessmodellen der Informationsverarbeitung.
3. **Empirische Untersuchung der Eigenschaften menschlicher kognitiver Prozesse:** Diese Methodologie stammt aus der Psychologie und dient neben den Ergebnissen der theoretischen Analysen als wesentliche Grundlage für den Entwurf von komputationalen Modellen.

Im Folgenden werden wir die Grundprinzipien der Computersimulation vorstellen. Von Seiten der Psychologie kann die Computersimulation als eine die Empirie ergänzende Forschungsmethode aufgefasst werden. Wie beim psychologischen Experiment (siehe Bortz 1984), können auch bei der Computersimulation zwei Vorgehensweisen unterschieden werden:

- **Hypothesengenerierend:** Ideen über mögliche Funktionsweisen menschlicher Informationsverarbeitung werden durch ein Computermodell beschrieben. Die Modellierung liefert dann den Ausgangspunkt für die Ableitung von (empirisch prüfbaren) Hypothesen.
- **Hypothesenprüfend:** Bestehende psychologische Theorien, die sich bereits zu einem gewissen Grad empirisch bewährt haben, können durch ihre Umsetzung in ein Computermodell präzisiert und logisch strukturiert werden. Zudem liefert die Umsetzung in eine Computersimulation den Nachweis der Konsistenz und Vollständigkeit der theoretischen Annahmen.

Für beide Aspekte der Computersimulation ist zu beachten, dass die Umsetzung von psychologischen Annahmen über Informationsverarbeitungsprozesse in ein Computerprogramm Freiheitsgrade enthält. Um ein Computerprogramm zu schreiben, das ein psychologisches Modell simuliert, müssen immer Zusatzannahmen getroffen werden (Cooper et al. 1996). Ein psychologisches Prozessmodell ist grundsätzlich immer durch verschiedene konkrete Computermodelle beschreibbar. Prinzipiell ist der Vergleich eines Computermodells mit dem informationsverarbeitenden System Mensch nur auf **funktionaler**, nicht aber auf **struktureller Ebene** möglich. D. h., das Computermodell verhält sich bei gegebenem Input ähnlich wie der Mensch; ob aber die interne Struktur, die zu diesem Verhalten führt, äquivalent ist, kann nicht entschieden werden (Searle 1980).

Als eine Methode zur Prüfung dieser funktionalen Äquivalenz wurde unter anderem der **Turing-Test** (Turing 1950) vorgeschlagen (■ Abb. 2.3). Dabei wird ein Beurteiler gebeten, über Computerterminals mit zwei Systemen zu kommunizieren, von denen ei-



■ **Abb. 2.3** Der Turing-Test illustriert eine Methode zur Bewertung der Fähigkeit eines Computers, menschliches Verhalten zu zeigen: Ein Beurteiler interagiert durch Fragen mit zwei anonymen Gesprächspartnern – einem Menschen und einem Computer, die beide nur durch eine Textkonsole auf jeweils einem Bildschirm kommunizieren. Der Beurteiler weiß nicht, welcher Bildschirm mit welchem Teilnehmer verbunden ist. Wenn der Computer so überzeugend agiert, dass ein beliebiger Beurteiler ihn nicht zuverlässig vom Menschen unterscheiden kann, gilt er als auf dem Niveau menschlicher Interaktion

nes ein Mensch ist, das andere ein Computerprogramm. Der Beurteiler weiß dabei nicht, welches der beiden Systeme der Mensch und welches das Computerprogramm ist. Beide Systeme reagieren über ihren jeweiligen Monitor auf die Eingaben des Beurteilers, welcher die Aufgabe hat, zu entscheiden, welche Antworten von einem Menschen getroffen werden. Kann der Beurteiler nicht entscheiden, welche Reaktionen vom Menschen und welche vom Computermodell gegeben werden, so hat das Computermodell den Test bestanden. Diese Art der Überprüfung kann jedoch zu völlig ungerechtfertigten Urteilen führen.

Besonders deutlich wird das an einem berühmten Gedankenexperiment von Searle (1980), dem sogenannten chinesischen Zimmer:

- » In einem Zimmer sitzt ein Mensch, der die chinesische Sprache nicht beherrscht. Er erhält Karten mit chinesischen Zeichen und eine Menge von Regeln, die festlegen, auf welche Zeichen er mit welchen anderen Zeichen reagieren soll. Ein Beurteiler, der die chinesische Sprache spricht, schiebt nun Karten mit chinesischen Schriftzeichen unter der Tür durch, der Mensch konsultiert seine Regeln und schiebt Karten zurück.

Der Beurteiler muss dabei zu dem Schluss kommen, dass der Mensch im Zimmer die chinesische Sprache beherrscht. Will man die Methode der Computersimulation sinnvoll zur Beschreibung menschlicher Informationsverarbeitung einsetzen, ist der Turing-Test nicht die angemessene Prüfmethode für die psychologische Adäquatheit eines Modells. Ein alternatives Vorgehen wäre folgendes: Ausgehend von vorliegenden Erkenntnissen über menschliche Gedächtnis- und Denkleistungen wird ein Prozessmodell für einen eingeschränkten Phänomenbereich entwickelt. Dieses Prozessmodell wird als Computermodell präzisiert. Dabei werden Kernannahmen identifiziert, die in empirisch prüfbare Hypothesen umgesetzt werden. Diese Hypothesen legen fest, wie sich bestimmte Zustände im kognitiven Prozess in beobachtbaren Größen, wie Fehler oder Reaktionszeiten, widerspiegeln. An die Stelle der globalen Zuschreibung von Eigenschaften, wie dies beim Turing-Test geschieht, tritt also eine detaillierte Analyse von kognitiven Prozessen.

## 2.4 Zur Vertiefung

In diesen Bereichen geben wir Ihnen Hinweise auf klassische Grundlagenartikel oder aktuelle Werke, in denen Sie mehr Informationen finden können.

### ■ ■ Einführung in Themen, Ansätze und Methoden der Kognitionsforschung

- Das folgende Buch ist ein Klassiker der Kognitionsforschung: Johnson-Laird, P. N. (1996). *Der Computer im Kopf: Formen und Verfahren der Erkenntnis*. dtv. Deutscher Taschenbuch Verlag.
- Ein aktueller Hintergrund zu menschlicher Kognition aus KI-Perspektive: Ragni, M. (2021). Kognition. In: Görz, G. and Schmid, U. and Braun, T., Hrsg. (eds.), *Handbuch der Künstlichen Intelligenz*, S. 227–278. De Gruyter Oldenbourg.

### ■ ■ Philosophischer Hintergrund

- In diesem Beitrag diskutiert Turing die Frage, ob Maschinen denken können: Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59, 433–460.

- In diesem Beitrag analysiert Searle die Frage von Turing im Kontext von Intentionalität und starker KI: Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.
- McCarthy, J. (1979). Ascribing Mental Qualities to Machines. Report no. ADA071423, Stanford University California, Dept. of Computer Science. Aufsatz, der unter anderem Searles Kritik provozierte.
- In diesem Beitrag stellt Newell die *Physical symbol system*-These vor, die maßgeblich für kognitive Systeme in einer „situated cognition“, also dass Kognition mit einer sozialen oder physischen Umwelt verbunden ist und durch den Kontext entsteht, geworden ist: Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4(2), 135–183.
- Dennett, D. C. (1989). *The Intentional Stance*. MIT Press, Cambridge, MA. Auseinandersetzung mit dem bei Dreyfus und Searle angesprochenen Problem, wann man menschlichen/künstlichen Systemen Intentionalität zuschreiben kann.
- Rationalität spielt in vielen Entscheidungen eine Rolle. Aus diesem Grund werden wir uns im ersten Teil des Buches auch mit normativen Aspekten wie der Logik als Grundlage für korrekte Schlussfolgerungen beschäftigen. Im zweiten Teil des Buches geht es dann um den deskriptiven Teil, also wie sich menschliches Denken (und damit auch die menschliche Rationalität) beschreiben lässt. In diesem umfangreichen Werk lässt sich der Hintergrund der aktuellen Rationalitätsdebatte nachlesen: Knauff, M. und Spohn, W. (2021). *The Handbook of Rationality*. MIT Press.



# Wissensrepräsentation und Logik

Wichtige Themen der KKI sind Wissensrepräsentation, Denken, Problemlösen, Lernen, Sprachverarbeitung, Bildverstehen und Expertise. Dabei ist die Frage der Repräsentation von Wissen für alle Bereiche grundlegend.

Ein zentraler Ansatz zur Repräsentation von Begriffen und Fakten ist die Struktur der semantischen Netze. Die Grundlagen für den Formalismus der semantischen Netze liefern die Mengentheorie und die Logik. Die folgenden Kapitel stellen diese Grundlagen schrittweise dar. In ► Kap. 3 wird die Mengenlehre eingeführt und gezeigt, wie hierarchische semantische Netze mit Konzepten der Mengenlehre formalisierbar sind. In ► Kap. 4 und 5 führen wir die Grundlagen der Aussagen- und der Prädikatenlogik ein und stellen die Formalisierung semantischer Netze mit Mitteln der Prädikatenlogik dar. In ► Kap. 6 führen wir die Methode des Theorembeweisens ein, mit der automatisch Schlussfolgerungen aus logisch repräsentiertem Wissen gezogen werden können. Teil I schließt ab mit einer Einführung in die Grundkonzepte der Programmiersprache Prolog ► Kap. 7. Wir zeigen, wie die in ► Kap. 3 eingeführten hierarchischen semantischen Netze in ein ablauffähiges Programm umgesetzt werden können.

## Inhaltsverzeichnis

**Kapitel 3      Grundlagen der Wissensrepräsentation – 19**

**Kapitel 4      Aussagenlogik – 33**

**Kapitel 5      Prädikatenlogik – 45**

**Kapitel 6      Schlussfolgern und Beweisen – 59**

**Kapitel 7      Logische Programmierung – 75**



# Grundlagen der Wissensrepräsentation

## Inhaltsverzeichnis

- 3.1      Begriffliches Wissen – 22
- 3.2      Strukturiertes Wissen: semantische Netze – 25
- 3.3      Zur Vertiefung – 32

Menschen *wissen* vieles über die Welt. Diese etwas plakative Aussage entspricht unserer Alltagserfahrung. Allerdings stellt sich im Zeitalter von fake news die Frage, was „Wissen“ eigentlich ist, wie es im Gedächtnis des einzelnen Menschen oder eines Computers repräsentiert wird und welche Wissensarten sich unterscheiden lassen.

### ► Beispiel 3.1

Angenommen, wir kommen an einem Spielplatz vorbei und hören, wie ein Mädchen sagt: „Es gibt in unserem Sonnensystem 11 Planeten.“ Offenbar ist dieser Satz objektiv betrachtet entweder wahr oder falsch. Er ist wahr, wenn unser Sonnensystem 11 Planeten hat, und falsch, wenn unser Sonnensystem mehr oder weniger als 11 Planeten hat. Wir wissen, dass die Astronomie aktuell nur 8 Himmelskörper als Planeten unseres Sonnensystems anerkennt. Dennoch könnte es aber sein, dass das Mädchen *überzeugt* davon ist, dass unser Sonnensystem 11 Planeten hat. In diesem Fall würden wir sagen, dass das Mädchen etwas Falsches glaubt, aber nicht, dass es etwas weiß, da Wissen in der Regel nur zutrifft, wenn das Geglaubte auch wahr ist. ◀

In der Informatik werden die Begriffe Überzeugung und Wissen nicht immer trennscharf verwendet. So wird manchmal mit Wissen (engl. *knowledge*) auch einfach die Menge abgespeicherter Daten bezeichnet, die dann natürlich wahr oder falsch sein können. In der Psychologie wird ebenfalls manchmal einfach von Wissen gesprochen und damit sowohl die objektive Wahrheit als auch die subjektive Überzeugung des individuellen Menschen gemeint.

### Definition 3.1 (Wissen und Überzeugung)

In der kognitiven Psychologie und der Künstlichen Intelligenz wird unter dem Begriff Wissen auch manchmal das bezeichnet, was eine Person oder ein Agent für *wahr hält*, also eine Überzeugung (engl. *belief*). (Strube 1993, S. 326)

Aus Gründen der Vereinfachung werden wir hier den Begriff des Wissens nutzen, auch wenn damit vor allem die individuellen Überzeugungen eines Menschen gemeint sind. Wie lässt sich Wissen nun im Sinne des Informationsverarbeitungsparadigmas verstehen (vgl. ► Kap. 2)? Die Grundlage dieses Paradigmas ist es, dass Informationsverarbeitungsprozesse als Transformationen von Symbolstrukturen beschreibbar sind. Diese Symbolstrukturen sind Repräsentationen von Überzeugungen im Gedächtnis, also Wissensrepräsentationen.

### Hintergrund

Die kognitive Psychologie unterscheidet unter dem Begriff „Wissen“ sowohl Faktenwissen (engl. *knowing that*) als auch prozedurales Wissen (engl. *knowing how*).

- **Faktenwissen**, oft auch als **deklaratives Wissen** bezeichnet, beinhaltet Aussagen, die wahr oder falsch sein können: „China liegt in Asien“, „Hunde sind Säugetiere“, „Die Währung von England ist das Pfund.“ Innerhalb des Faktenwissens wird zwischen **semantischem** (begrifflichem) und **episodischem** Wissen unterschieden (vgl. Tulving 1972). Die genannten Beispiele gehören zum semantischen Wissen. Unter episodischem Wissen wird Wissen über individuelle Erfahrungen verstanden, also zum Beispiel, wenn Hannah weiß, dass sie gestern auf Netflix „Star Trek“ gesehen hat. Faktenwissen kann auch **unscharf** sein. So wissen wir beispielsweise, dass es Anfang November meistens regnet oder dass Wohnungen in Berlin kaum unter €300 gemietet werden können.
- **Prozedurales Wissen** umfasst Handlungswissen, also das Wissen, wie ich ein konkretes Problem lösen kann, z. B. „Wenn ich Tee kochen will, muss ich Wasser heiß machen“ oder „Wenn ich ein Problem nicht abstrakt lösen kann, versuche ich es zunächst an einem konkreten Beispiel.“ Außerdem wird dazu auch heuristisches Wissen gezählt, welches nicht konkret, sondern allgemein und flexibel ist. Heuristisches Wissen

umfasst grobe Faustregeln oder intuitive Strategien, die auf verschiedene Situationen angewandt werden können, wie zum Beispiel die Möglichkeit, den kürzesten Weg von *A* nach *B* durch die Nähe von Wegen zur Luftlinie zwischen *A* und *B* abzuschätzen.

Um Wissen so zu repräsentieren, dass es von einem Computer verarbeitet werden kann, wurden verschiedene Ansätze der **Wissensrepräsentation** vorgeschlagen. Als Methoden der Wissensrepräsentation werden meistens Logik, hierarchische und nichthierarchische semantische Netze, Schemata, Skripts, Produktionsregeln und analoge Repräsentationen angeführt (Barr et al. 1981, Kap. 3). Um diese verschiedenen Repräsentationsformalismen zu vergleichen, kann man verschiedene Kriterien verwenden, wie wir im Folgenden aufzeigen.

Ein für die Psychologie relevantes Kriterium ist das der psychologischen Adäquatheit, d. h. dass die gewählte beschreibende Repräsentation sich so verhält wie die tatsächliche mentale Repräsentation. Es ist jedoch kaum möglich, Repräsentationsannahmen so zu operationalisieren, dass Beweise für die Adäquatheit eines Repräsentationsformalismus gewonnen werden können. Eine abgeschwächte Möglichkeit zur Prüfung der Adäquatheit wäre die psychologische Plausibilität, d. h., ist der Repräsentationsformalismus einfach kognitiv verarbeitbar und hilft die Repräsentation, Informationsverarbeitungsprozesse vorherzusagen (Larkin und Simon 1987)? Drei allgegenwärtige Prozesse der Wissensverarbeitung (s. ► Kap. 1) sind:

- der Erwerb neuen Wissens,
- der Abruf von Wissen,
- das Schlussfolgern und Problemlösen auf der Grundlage von Wissen.

In diesem Kapitel beschäftigen wir uns mit der Repräsentation von explizit gegebenem Wissen. Auf Mechanismen (► Kap. 4, Logik), wie sich aus implizitem Wissen zum Beispiel durch Schlussfolgerungen explizites Wissen generieren lässt, wird im nächsten Kapitel eingegangen.

Im Folgenden wird ein kurzer Überblick über verschiedene Repräsentationsformate gegeben. Die **Logik** ist ein grundlegendes Werkzeug zum Schlussfolgern: D. h., wenn eine bestimmte Menge von Fakten gegeben ist, können mithilfe logischer Schlussregeln Folgerungen aus diesen Fakten gezogen werden. So kann etwa aus den Fakten „Alle Menschen sind sterblich“ und „Sokrates ist ein Mensch“ geschlossen werden, dass Sokrates sterblich ist. Die verwendete Schlussregel ist ein **Syllogismus**. Die genannten Fakten bilden die *Prämissen* des Syllogismus, der daraus abgeleitete Fakt heißt *Konklusion*. Umgangssprachlich können wir den Schluss folgendermaßen formulieren: „WENN alle Menschen sterblich sind UND Sokrates ein Mensch ist, DANN ist Sokrates sterblich.“ In den Symbolen der **Prädikatenlogik** (► Kap. 5) kann der Schluss folgendermaßen dargestellt werden:

$$\begin{aligned} \forall x (\text{mensch}(x) \rightarrow \text{sterblich}(x)) \\ \text{mensch}(\text{sokrates}) \\ \text{sterblich}(\text{sokrates}) \end{aligned}$$

Ein rein logischer Kalkül liefert keine Möglichkeit, Wissen strukturiert zu repräsentieren. Eine solche Möglichkeit bilden semantische Netze und häufig verwendete Spezialformen wie hierarchische Netze und das Schemakonzept (engl. *frame*), auf die in ► Abschn. 3.2 eingegangen wird.

### 3.1 Begriffliches Wissen

Ein Aspekt menschlichen Wissens ist, dass es in Begriffen strukturiert ist. Begriffliches Wissen kann hierarchisch organisiert sein. So wissen wir beispielsweise, dass Hunde und Katzen Säugetiere sind, dass Säugetiere Tiere sind und dass sich Tiere bewegen können. Wir fassen Sorten von Dingen zu umfassenderen Sorten. Wir nennen das „eine Menge bilden“. Solche Mengen von Dingen lassen sich wiederum zu größeren Mengen zusammenfassen. Die Menge aller Katzen etwa ist eine Teilmenge aller Säugetiere. Mengen und Operationen auf Mengen bilden also eine wichtige Grundlage zur Beschreibung und Kategorisierung unserer Ontologie, d. h. der Dinge, die wir annehmen.


#### Definition 3.2 (Menge)

Eine **Menge** ist eine Gesamtheit von Dingen oder abstrakten Einheiten, die wir als **Elemente** dieser Menge bezeichnen. Ist ein Objekt  $e$  Element der Menge  $\mathbb{M}$ , so schreiben wir  $e \in \mathbb{M}$ ; ist  $e$  kein Element von  $\mathbb{M}$ , so schreiben wir  $e \notin \mathbb{M}$ . Eine Menge ist gerade durch ihre Elemente bestimmt: Wenn  $\mathbb{M}$  und  $\mathbb{N}$  genau dieselben Elemente haben, dann ist  $\mathbb{M} = \mathbb{N}$ .

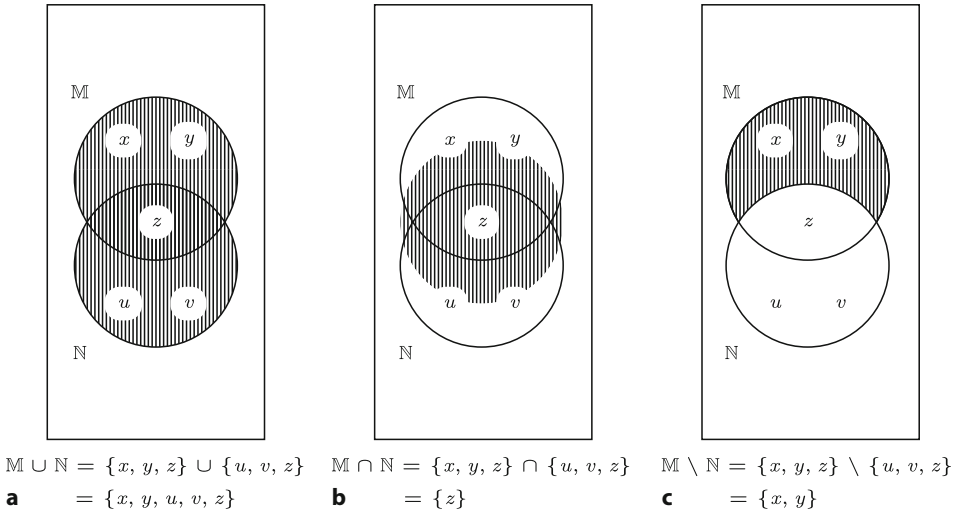
So ist zum Beispiel  $\mathbb{M} = \{\text{Stuhl, Schrank, Tisch}\}$  eine Menge. Es gilt  $\text{Stuhl} \in \mathbb{M}$  und  $\text{Palme} \notin \mathbb{M}$ . Bei der Darstellung einer Menge ist die Reihenfolge der Elemente beliebig, auch mögliche Wiederholungen von Elementen ändern nichts an der Menge. Es gilt also

$$\{\text{Stuhl, Schrank, Tisch}\} = \{\text{Tisch, Schrank, Stuhl}\} = \{\text{Tisch, Schrank, Tisch, Stuhl}\}.$$

Es ist Konvention, die jeweils sparsamste Schreibweise der Menge zu notieren, hier also zum Beispiel  $\{\text{Tisch, Schrank, Stuhl}\}$  oder  $\{\text{Stuhl, Tisch, Schrank}\}$ . Eine einelementige Menge ist nicht dasselbe wie ihr eines Element. Es gilt  $\text{Stuhl} \in \{\text{Stuhl}\}$ , aber nicht  $\text{Stuhl} = \{\text{Stuhl}\}$ . Mengen können ineinander verschachtelt sein. Es gibt also Mengen, deren Elemente wieder Mengen sind. Die Menge  $\mathbb{M} = \{\{a, b\}\}$  besitzt genau ein Element, nämlich die Menge  $\{a, b\}$ . Es gilt:  $\{a, b\} \in \mathbb{M}$  und  $a \notin \mathbb{M}$ ,  $b \notin \mathbb{M}$ . Wichtige Operationen auf Mengen sind die Vereinigung ( $\cup$ ), der Schnitt ( $\cap$ ) und die Mengendifferenz ( $\setminus$ ) zweier Mengen.

**Vereinigung zweier Mengen** Die Vereinigung von zwei Mengen  $\mathbb{M}$  und  $\mathbb{N}$ , geschrieben  $\mathbb{M} \cup \mathbb{N}$ , ist die Menge aller Elemente, die zu  $\mathbb{M}$  oder  $\mathbb{N}$  gehören, also  $\mathbb{M} \cup \mathbb{N} = \{e \mid e \in \mathbb{M} \text{ oder } e \in \mathbb{N}\}$ . So können zum Beispiel die Menge der Küchenmöbel und die Menge der Wohnzimmermöbel zu einer Menge vereinigt werden:  $\text{Möbel} = \text{Küchenmöbel} \cup \text{Wohnzimmermöbel}$ . Das ist natürlich nur vereinfachend gemeint; Schlafzimmermöbel sind natürlich auch Möbel usw. Wichtig bei der Vereinigung zweier Mengen ist, dass Elemente, die in beiden Mengen vorkommen, in der Vereinigungsmenge nur einmal vorhanden sind. Wenn also  $\mathbb{A} = \{a, b, c, d\}$  und  $\mathbb{B} = \{c, d, e, f\}$ , dann ist  $\mathbb{A} \cup \mathbb{B} = \{a, b, c, d, e, f\}$ . Eine grafische Veranschaulichung der Vereinigung gibt  Abb. 3.1a.

**Schnitt zweier Mengen** Der Schnitt oder Durchschnitt zweier Mengen  $\mathbb{M}$  und  $\mathbb{N}$ , geschrieben  $\mathbb{M} \cap \mathbb{N}$  (die Schnittmenge), ist die Menge aller Elemente, die sowohl zu  $\mathbb{M}$  als auch zu  $\mathbb{N}$  gehören, also  $\mathbb{M} \cap \mathbb{N} = \{e \mid e \in \mathbb{M} \text{ und } e \in \mathbb{N}\}$ . Zwei Mengen, die keine gemeinsamen Elemente besitzen, heißen disjunkt. Zwei Mengen sind also disjunkt,



■ **Abb. 3.1** Mengen und die Operationen darauf lassen sich durch Venn-Diagramme repräsentieren. So gehören oben zur Menge  $M$  die Elemente  $x, y, z$  und zur Menge  $N$  die Elemente  $u, v, z$ . Werden die beiden Mengen vereinigt zu  $\mathbb{M} \cup \mathbb{N}$ , erhält man die schraffierte Menge in Abbildung (a), beim Schnitt  $\mathbb{M} \cap \mathbb{N}$  die schraffierte Menge in Abbildung (b) und bei der Differenzbildung  $\mathbb{M} \setminus \mathbb{N}$  die schraffierte Menge in Abbildung (c)

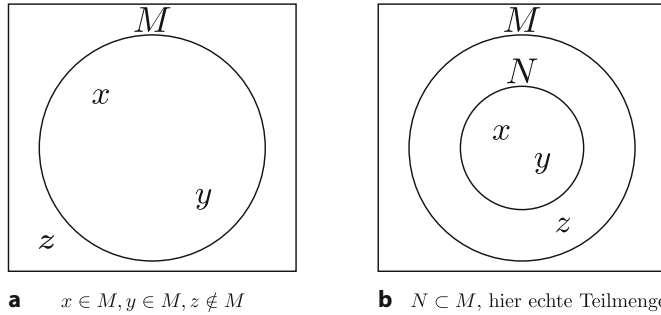
wenn  $\mathbb{M} \cap \mathbb{N} = \emptyset$  ist (wobei  $\emptyset$  die leere Menge ist: die Menge, die überhaupt keine Elemente hat). Für  $\mathbb{A} = \{a, b, c, d\}$  und  $\mathbb{B} = \{c, d, e, f\}$  ist  $\mathbb{A} \cap \mathbb{B} = \{c, d\}$ . Eine grafische Veranschaulichung der Schnitt-Operation gibt ■ Abb. 3.1b.

**Die Mengendifferenz** Das **absolute Komplement** einer Menge meint alle Elemente, die außerhalb dieser Menge liegen, und wird mit  $\setminus \mathbb{A}$  notiert. Das absolute Komplement kann man dann angeben, wenn man nicht nur weiß, was zu einer Menge gehört, sondern auch weiß, was es alles außerhalb dieser Menge gibt. Ist unsere Gesamtmenge beispielsweise die Menge aller Tiere und  $\mathbb{A}$  die Menge aller Katzen, so ist  $\setminus \mathbb{A}$  die Menge aller Tiere, die *keine* Katzen sind. Das **relative Komplement** zweier Mengen  $\mathbb{A}$  und  $\mathbb{B}$ , auch Differenzmenge genannt, ist die Menge aller Elemente, die zu  $\mathbb{A}$ , aber nicht zu  $\mathbb{B}$  gehören, und wird als  $\mathbb{A} \setminus \mathbb{B}$  oder  $\mathbb{A} - \mathbb{B}$  notiert (lies „ $\mathbb{A}$  ohne  $\mathbb{B}$ “). Eine grafische Veranschaulichung der Differenzmengenoperation gibt ■ Abb. 3.1c.

Die Zugehörigkeit zu einer Menge und Mengenoperationen können grafisch in einem sogenannten **Venn-Diagramm** veranschaulicht werden (siehe ■ Abb. 3.1).

Eine Menge  $\mathbb{A}$  kann Teil einer anderen Menge  $\mathbb{B}$  sein, nicht in dem Sinne, dass  $\mathbb{A}$  ein Element von  $\mathbb{B}$  wäre, sondern weil alle Elemente von  $\mathbb{A}$  auch Elemente von  $\mathbb{B}$  sind. Diese Beziehung heißt Inklusion oder Teilmengenbeziehung und wird mit  $\subset$  oder  $\subseteq$  notiert. Dabei heißt  $\subsetneq$  „**echte Teilmenge**“;  $\mathbb{A} \subset \mathbb{B}$  bedeutet, dass  $\mathbb{A}$  nur einen Teil der Elemente von  $\mathbb{B}$  umfasst (es gibt mindestens ein Element von  $\mathbb{B}$ , das nicht Element von  $\mathbb{A}$  ist). Bei  $\mathbb{A} \subseteq \mathbb{B}$  kann  $\mathbb{A}$  auch alle Elemente von  $\mathbb{B}$  umfassen. In diesem Fall gilt  $\mathbb{A} \subseteq \mathbb{B}$  und  $\mathbb{B} \subseteq \mathbb{A}$ , und es folgt  $\mathbb{A} = \mathbb{B}$ . Für die leere Menge gilt:  $\emptyset \subseteq \mathbb{M}$ , egal was die Menge  $\mathbb{M}$  ist.

Die Menge aller Möbel kann zum Beispiel die Küchenmöbel als Teilmenge enthalten: Küchenmöbel  $\subseteq$  Möbel. Eine grafische Veranschaulichung der Teilmengenbeziehung ist in ■ Abb. 3.2b gegeben.



■ **Abb. 3.2** Venn-Diagramme für (a) Menge–Element-Beziehung und (b) Teilmengenbeziehung

### Hintergrund: Endliche und unendliche Mengen

Mengen können unterschiedlich viele Elemente haben. So kann man die Menge der Teilnehmer eines Seminars alle aufzählen, während man die Menge der natürlichen Zahlen  $\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}$  nicht vollständig aufschreiben kann. Die Menge der Teilnehmer des Seminars ist endlich, die Menge der natürlichen Zahlen unendlich. Bei den unendlichen Mengen wird noch unterschieden, ob sie **abzählbar** oder **überabzählbar** unendlich sind. Die Menge der natürlichen Zahlen kann durch die Vorschrift „addiere 1 zum letzten Element“ konstruiert werden; sie ist abzählbar unendlich. Bei den reellen Zahlen ist es nicht möglich, eine solche Konstruktionsvorschrift für aufeinanderfolgende Zahlen anzugeben. Zwischen zwei Elementen (zum Beispiel 0 und 0,1) liegen jeweils wieder unendlich viele Elemente (zum Beispiel 0,01, 0,001, ...). Die Menge der reellen Zahlen ist überabzählbar unendlich, das bedeutet, dass diese echt größer ist als die Menge der natürlichen oder der rationalen Zahlen. Unendliche Mengen können durch Angabe eines Teils ihrer Elemente, zum Beispiel  $\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}$ , dargestellt werden. Dies ist aber nur sinnvoll, wenn man annehmen kann, dass ein Betrachter das Bildungsgesetz der Menge versteht. Ansonsten gibt man das Bildungsgesetz der Menge besser explizit an. Allgemein wird ein Mengenbildungsgesetz folgendermaßen angegeben:  $\mathbb{M} = \{e \mid E(e)\}$ , d. h., „ $\mathbb{M}$  ist die Menge aller  $e$ , die die Eigenschaft  $E$  haben.“ Zum Beispiel: Frauen =  $\{e \mid e \text{ ist weiblich}\}$ . Es kann sein, dass man Bildungsgesetze definiert, für die es keine Objekte gibt. Dann ist die definierte Menge leer. Die leere Menge wird als  $\{\}$  oder  $\emptyset$  notiert. So ist zum Beispiel  $\mathbb{X} = \{x \mid x \text{ ist Säugetier und } x \text{ hat Kiemen}\}$  die leere Menge.

Wie wir gesehen haben, kommt es bei Mengen nicht darauf an, in welcher Reihenfolge man die Elemente notiert und ob man dabei manche wiederholt. In der Künstlichen Intelligenz gibt es zwei Konzepte, die dieses Prinzip ergänzen: Listen und Tupel. Bei **Listen**<sup>1</sup> ist die Reihenfolge der Elemente hingegen relevant. Elemente, die mehrfach in der Liste vorkommen, bleiben erhalten, auch wenn sie denselben Namen haben. Im Gegensatz zu Mengen werden Listen mit runden Klammern notiert. Es gilt also (Stuhl, Schrank, Tisch)  $\neq$  (Tisch, Schrank, Stuhl)  $\neq$  (Tisch, Schrank, Tisch, Stuhl, Stuhl). Listen mit einer fest vorgegebenen Anzahl von Elementen werden als **Tupel** bezeichnet. 2-Tupel, also Tupel mit genau zwei Komponenten, werden auch „geordnete Paare“ genannt, und 3-Tupel nennt man „Tripel“. Der Unterschied zwischen Tupeln und Listen besteht darin, dass Tupel eine feste Anzahl von Elementen haben und unveränderlich sind, während Listen veränderlich sind und ihre Länge dynamisch angepasst werden kann.

1 Eine Einkaufsliste ist ein anschauliches Beispiel für eine Liste: Die Reihenfolge der Elemente kann eine Rolle spielen, beispielsweise wenn sie in der Reihenfolge der Geschäfte angeordnet ist, in denen man etwas besorgen möchte.

### Hintergrund: Rechengesetze für Mengen

Die eingeführten Grundoperationen ermöglichen es, mit Mengen zu rechnen. Die Rechengesetze, die im Folgenden vorgestellt werden, können mit Venn-Diagrammen nachgewiesen werden, was man zur Übung tun kann. Für Mengen  $A, B, M$  gilt:

1. Wenn  $A \subseteq B$  und  $B \subseteq M$ , dann  $A \subseteq M$ . (Transitivität der Teilmengenbeziehung)
2.  $A \cup B = B \cup A$  und  $A \cap B = B \cap A$ . (Kommutativität von Vereinigung und Schnitt)
3.  $(A \cup B) \cup M = A \cup (B \cup M)$  und  $(A \cap B) \cap M = A \cap (B \cap M)$ . (Assoziativität von Vereinigung und Schnitt)
4.  $A \cup \emptyset = A$  und  $A \cap \emptyset = \emptyset$ . (neutrales Element)
5.  $(A \cup B) \cap M = (A \cap M) \cup (B \cap M)$  und  $(A \cap B) \cup M = (A \cup M) \cap (B \cup M)$ . (Distributivität)
6.  $A \subseteq A \cup B$  und  $B \subseteq A \cup B$ . Wenn  $A \subseteq M$  und  $B \subseteq M$ , dann  $A \cup B \subseteq M$ .
7.  $A \cap B \subseteq A$  und  $A \cap B \subseteq B$ . Wenn  $M \subseteq A$  und  $M \subseteq B$ , dann  $M \subseteq A \cap B$ .
8.  $M \setminus (A \cup B) = (M \setminus A) \cap (M \setminus B)$  und  $M \setminus (A \cap B) = (M \setminus A) \cup (M \setminus B)$ . (Regeln von De Morgan)

## 3.2 Strukturiertes Wissen: semantische Netze

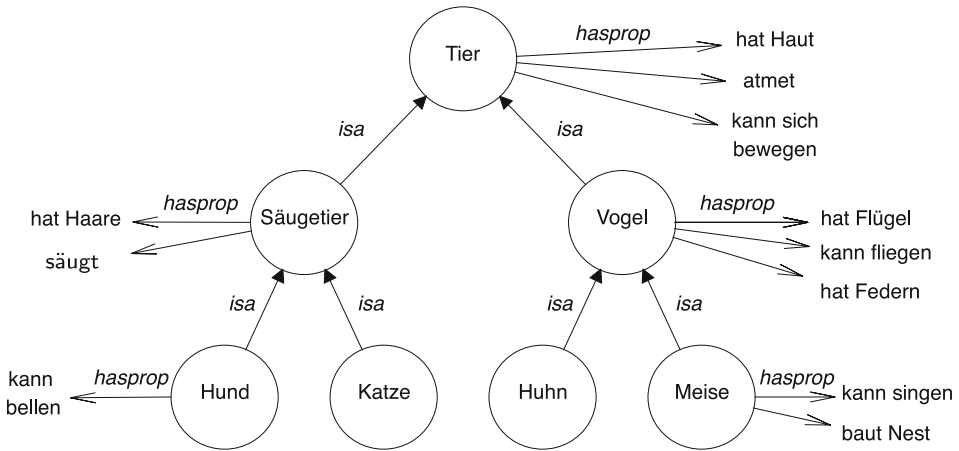
Es ist psychologisch plausibel (und in Experimenten zu semantischen Netzen und im Rahmen des Schemaansatzes immer wieder belegt, z. B. Wender 1988), dass Menschen Wissen in strukturierter Form repräsentieren, ihr Wissen also organisiert ist. Zur organisierten Repräsentation von begrifflichem Wissen werden semantische Netze verwendet. Ein semantisches Netz besteht allgemein aus „Knoten“ und „Kanten“, wobei eine Kante immer zwei (eventuell denselben) Knoten verbindet. Eine solche Struktur wird als **(gerichteter) Graph** bezeichnet (► Abschn. 9.1). Knoten repräsentieren dabei Begriffe und Eigenschaften, Kanten geben Beziehungen zwischen zwei Knoten an. In hierarchischen semantischen Netzen, mit denen wir uns im Folgenden beschäftigen wollen, existieren zwei Arten von Kanten:

- Kanten, die eine Beziehung zwischen Begriffen angeben. Diese werden durch *isa*-Kanten repräsentiert; der Name ist abgeleitet vom Englischen *is a* („ist ein“).
- Kanten, die einem Begriff eine Eigenschaft zuordnen. Diese werden durch sogenannte *hasprop*-Kanten repräsentiert; der Name ist vom Englischen *has property* („hat [die] Eigenschaft“) abgeleitet.

Zur Veranschaulichung zeigt ■ Abb. 3.3 einen Ausschnitt des Netzwerkmodells von Collins und Quillian (1969). In diesem Netz wird die hierarchische Struktur eines Ausschnitts unseres begrifflichen Wissens repräsentiert. Die einem Begriff zugeordneten Eigenschaften gelten auch für alle dem Begriff untergeordneten Knoten und müssen damit nur einmal gespeichert werden.

Um für einen Begriff zu entscheiden, ob ihm eine bestimmte Eigenschaft zukommt, wird ausgehend von diesem Begriff (Knoten) gesucht, ob er durch *hasprop*-Kanten direkt oder indirekt mit dem Knoten dieser Eigenschaft verbunden ist. Soll etwa geprüft werden, ob ein Hund atmet, wird ausgehend von dem Knoten, der den Begriff „Hund“ repräsentiert, zunächst die mit diesem Knoten verbundene *hasprop*-Kante („kann bellen“) betrachtet. Wird die Eigenschaft dort nicht gefunden, so wird eine *isa*-Ebene höher, am Säugetier-Knoten, weiter gesucht. Die Eigenschaft „atmet“ ist dem Begriff Tier zugeordnet, dem Säugetiere und damit auch Hunde untergeordnet sind. Verfolgt man also den Weg vom Knoten „Hund“ bis zum Knoten „Tier“ entlang der *isa*-Kanten, so kann die Frage, ob ein Hund atmet, positiv beantwortet werden. Collins und Quillian konnten in Experimenten zumindest teilweise belegen, dass die Antwortzeiten für die Entscheidung, ob eine Eigenschaft für einen Begriff gilt, um so länger werden, je mehr *isa*-Kanten zwi-





■ **Abb. 3.3** Ein semantisches Netzwerk über Tierarten, mit einigen zugehörigen Eigenschaften (repräsentiert durch *hasprop*)

schen dem entsprechenden Begriff und dem Oberbegriff, dem die gefragte Eigenschaft zugeordnet ist, liegen (z. B. Anderson 2020, S. 127).

Betrachten wir nun zunächst die *isa*-Kanten genauer. Bei der *isa*-Relation wird zusätzlich zwischen Teilmengenbeziehungen (**Inklusionen**) und Zugehörigkeit (**Instanzen**) unterschieden. Bei der Inklusion handelt es sich um eine Beziehung zwischen zwei Mengen:  $\text{Hunde} \subseteq \text{Säugetiere}$ . Bei der Instanziierung handelt es sich dagegen darum, dass ein Objekt Element einer Menge ist (bzw. unter den entsprechenden Begriff fällt):  $\text{meinHundTobi} \in \text{Hunde}$ . Zur Vereinfachung werden im Folgenden nur die Teilmengenbeziehungen weiter betrachtet. Die Inklusion ist eine **Relation**, genauer: eine **Ordnungsrelation**. Diese mathematischen Begriffe sollen zunächst erläutert werden.

### Definition 3.3 (Mengenrelation)

Das kartesische Produkt zweier Mengen  $\mathbb{M}$  und  $\mathbb{N}$  ist definiert als die Menge aller geordneten Paare  $(x, y)$  mit  $x \in \mathbb{M}$  und  $y \in \mathbb{N}$ , also  $\mathbb{M} \times \mathbb{N} = \{(x, y) \mid x \in \mathbb{M} \text{ und } y \in \mathbb{N}\}$ . Dabei kann auch  $\mathbb{M} = \mathbb{N}$  sein. Eine **Relation**  $\mathcal{R}$  zwischen den Mengen  $\mathbb{M}$  und  $\mathbb{N}$  ist eine Teilmenge des kartesischen Produkts  $\mathbb{M} \times \mathbb{N}$ , also  $\mathcal{R} \subseteq \mathbb{M} \times \mathbb{N}$ .

Relationen sind spezielle Mengen, nämlich solche, deren Elemente geordnete Paare sind. In der obigen Definition wurde eine Relation über zwei Mengen definiert. Dies ist ein Spezialfall einer Relation, eine sogenannte zweistellige Relation. Allgemein können Relationen über eine beliebige Anzahl von Mengen definiert werden. Eine  $n$ -stellige Relation ist dann eine Menge von  $n$ -Tupeln. Zum Beispiel ist

$$\mathcal{R} \subseteq \mathbb{L} \times \mathbb{M} \times \mathbb{N} = \{(x, y, z) \mid x \in \mathbb{L} \text{ und } y \in \mathbb{M} \text{ und } z \in \mathbb{N}\}$$

eine dreistellige Relation, also eine Menge von 3-Tupeln. Zur Veranschaulichung kann man sich eine Datenbank eines Arbeitgebers in der Bauwirtschaft vorstellen, in der es eine Menge von Namen, spezifischen Berufsbezeichnungen und Arbeitsorten gibt. Eine Relation wären dann alle Maurer, die bei diesem Unternehmen beschäftigt sind. Auch **Funktionen**, wie zum Beispiel  $f(x) = x^2$ , können als Mengen von Tupeln, also als Relationen,