

# Kapitel 1

## Symbole, Objekte und Konzepte

Um die potenzielle Bedeutung von Ontologien in der Informationsverarbeitung zu verstehen, muss man sich klarmachen, was die grundlegenden Aufgaben und Probleme der Informationsverarbeitung sind: Nämlich bestimmte Ausschnitte der realen Welt in eine geeignete Darstellungsform zu überführen, die mit Hilfe des Computers manipuliert werden kann, um mit diesen Veränderungen in der realen Welt abzubilden. Dies können wir am Beispiel des Electronic Banking genauer betrachten. Ein System in diesem Bereich muss bestimmte Objekte der Welt – Personen, Organisationen, Konten, Aktien, Geldbeträge usw. – abbilden. Diese Objekte können physikalische Objekte sein, wie etwa Personen, oder aber abstrakte Objekte wie Konten. Häufig ist hierbei der Übergang zwischen konkreten und abstrakten Objekten fließend. Eine Aktie zum Beispiel kann ein physikalisches Objekt, nämlich ein Anteilschein in Form eines Dokuments, oder aber ein abstrakter Anteil an einem Unternehmen sein. Operationen, die auf den im System enthaltenen Repräsentationen von Personen, Konten usw. ausgeführt werden, spiegeln hierbei Veränderungen der realen Welt wider. Die

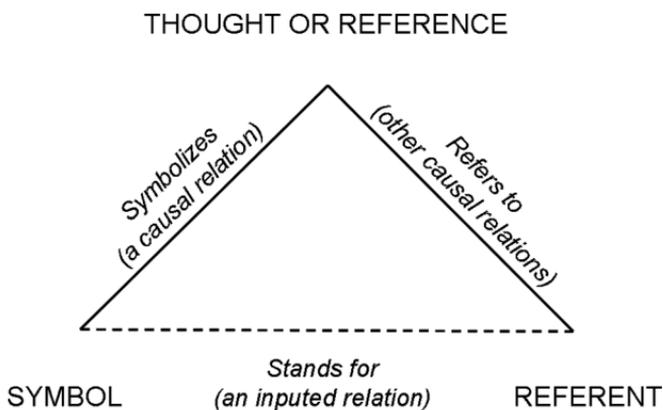
Überweisung eines Geldbetrags dargestellt durch die Subtraktion eines bestimmten Betrags von einer Zahl im System, sowie die Addition des gleichen Betrags auf eine andere Zahl<sup>1</sup>, hat hierbei eine konkrete Veränderung der Besitzverhältnisse in der realen Welt zur Folge.

Da, wie das Beispiel zeigt, Informationsverarbeitung unmittelbaren Einfluss auf die reale Welt haben kann, ist die korrekte und angemessene Abbildung der realen Welt in das entsprechende Programm ein zentrales Problem. Die Definition einer solchen Abbildung wird hierbei dadurch erschwert, dass es in der Regel eine Vielzahl unterschiedlicher Möglichkeiten gibt, diese Abbildung vorzunehmen. Welche Darstellung die 'beste' ist, lässt sich oft nicht sagen und hängt von unterschiedlichen Faktoren ab: dem Ausschnitt der betrachteten Welt, der Verfügbarkeit von Informationen, die zu erfüllende Aufgabe, die technischen Möglichkeiten, usw. Diese Fülle an Möglichkeiten führt dazu, dass Systeme, die ganz ähnliche Aufgaben erfüllen, sehr unterschiedliche Darstellungen der realen Welt verwenden. Unterschiede zwischen verwendeten Darstellungen können rein struktureller Natur sein, wie etwa die getrennte Speicherung von Vor- und Nachnamen versus der gemeinsamen Speicherung in einem einzigen Datenfeld. Mögliche Unterschiede können aber auch viel grundlegenderer Natur sein. Eine Frage ist zum Beispiel, ob Organisationen und Personen als Inhaber von Konten einheitlich oder in unterschiedlichen Repräsentationen dargestellt werden. Erfolgt eine einheitliche Darstellung schließt sich die Frage an, ob Namen von Organisationen und Personen gleich behandelt werden können (was ist der Vorname eines Unternehmens?) und was die Gemeinsamkeiten von Personen und Organisationen sind, die eine solche Gleichbehandlung rechtfertigen. Aus rechtlicher Sicht sind beide juristische Personen. Dennoch bestehen Unterschiede in der

---

<sup>1</sup> Der Vorgang bei einer Umbuchung ist natürlich viel komplexer, was aber an dieser Stelle nichts zur Sache tut.

Behandlung von Privat- und Firmenkunden, die sich nicht nur aus strategischen Überlegungen der entsprechenden Bank, sondern auch aus gesetzlichen Vorgaben ergeben (Stichwort Beratungspflicht). Schon an diesem einfachen Beispiel zeigen sich die Schwierigkeiten, denen sich die Informationsverarbeitung tagtäglich stellen muss. Hinzu kommt der Trend der Vernetzung von Informationen - Daten über Konten werden nicht nur bei der entsprechenden Bank sondern auch bei der Schufa und neuerdings von diversen Behörden geführt und verwendet. Verwenden die Systeme Darstellungen, die von denen der entsprechenden Bank abweichen, wird der Austausch von Informationen erschwert und erfordert ein genaueres Verständnis der jeweils gewählten Darstellungsformen.



**Abb. 1.1** Das Semiotische Dreieck

Wir können demnach feststellen, dass die Abbildung realer Objekte auf eine entsprechende Darstellung auf der damit verbundenen Erwartungen bzgl. der Natur der dargestellten Objekte beruht. Dieser komplexe Zusammenhang wurde bereits Anfang des Jahrhunderts im Bereich der Semiotik, der Lehre von

Zeichen und ihrer Bedeutung untersucht. Entsprechende Überlegungen finden sich bereits bei Aristoteles, eine explizite Beschreibung der Beziehung zwischen einer Darstellung (in Form eines Symbols), dem realen Objekt und den damit verbundenen Erwartungen findet sich bei Odgen und Richardson in Form des sogenannten semiotischen Dreiecks (vgl. Abbildung 1.1). Ausgangspunkt der Betrachtung sind hierbei Symbole in einer bestimmten Darstellungsform. Diese Symbole beziehen sich auf Objekte in der realen Welt, die in der Abbildung als Referenten bezeichnet werden. In unserem Beispiel wäre die Kundennummer ein solches Symbol, welches sich auf den entsprechenden Kunden, sei es eine Person oder eine Organisation bezieht. Als drittes Element führen Odgen und Richardson die Erwartungen ein, die der Betrachter mit dem Objekt bzw. dem Symbol verbindet. Diese werden hier als 'Thought or Reference' bezeichnet. In neueren Arbeiten wird diese Bezeichnung häufig durch den Begriff des Konzeptes ersetzt, eine Bezeichnung, die auch im Bereich der Informationsverarbeitung gebräuchlich ist, um Klassen von Objekten und deren typische Eigenschaften zu beschreiben. Die Rolle der Ontologien in der Informationsverarbeitung ist nun, grob gesagt, die Formalisierung dieses dritten Elements, welches die Verbindung zwischen Symbolen einer informationstechnischen Darstellung und den Erwartungen an das hierdurch dargestellte Objekt herstellt. Um diese Rolle genauer zu verstehen, ist eine genauere Betrachtung dieses dritten Elements sowie dessen Relationen zu Objekten auf der einen, und Symbolen auf der anderen Seite erforderlich.

## 1.1 Konzepte und das Universalienproblem

Um die Relation zwischen Symbolen, Objekten der realen Welt und Konzepten zu verstehen, ist ein kleiner Exkurs in den Bereich der philosophischen Ontologie notwendig. Die Untersu-

chung der Natur von Konzepten und deren Rolle bei der Beschreibung der Welt ist eine zentrale Fragestellung der Ontologie, häufig als 'Universalienproblem' bezeichnet. Um dieses Problem zu veranschaulichen betrachten wir natürlichsprachliche Aussagen wie die folgenden:

"Die Parkbank ist grün."

"Herr Meier ist ein Kunde."

Versucht man einen Bezug zwischen diesen Sätzen und der realen Welt herzustellen, um deren Bedeutung zu erfassen, so stellt man fest, dass dies nicht ohne Weiteres möglich ist. Die Interpretation der Satzteile 'Herr Meier' sowie 'Die Parkbank' ist relativ einfach, da wir diese auf konkrete Objekte in der Welt, nämlich die entsprechende Person bzw. das entsprechende Sitzmöbel abbilden können. Auch die Interpretation der Sätze als Ganzes ist leicht möglich, indem diesen ein Wahrheitswert zugeordnet wird. Wir betrachten die reale Welt und entscheiden, ob die entsprechende Aussage wahr ist, also ob die bezeichnete Bank tatsächlich grün bzw. Herr Meier tatsächlich ein Kunde ist. Schwieriger gestaltet sich die korrekte Abbildung der Eigenschaften der Parkbank bzw. der Eigenschaften von Herrn Meier in die reale Welt. In beiden Fällen gibt es kein konkretes Objekt, welches Ziel der Abbildung ist, da es in der realen Welt weder das Objekt 'grün' noch das Objekt 'Kunde' gibt. Es handelt sich hierbei vielmehr um eine allgemeine Eigenschaft, die verschiedene Objekte gemeinsam haben können. Solche gemeinsamen Eigenschaften werden auch Universalien genannt, da sie kein Individuum, sondern eine universellere Beschreibung darstellen. Konkrete Objekte können hierbei als Beispiele für diese universellen Beschreibungen dienen, etwa Herr Meier als Beispiel für einen Kunden. Während es unstrittig ist, dass universelle Beschreibungen in der Sprache ihren festen Platz haben, herrscht in der Philosophie Uneinigkeit darüber, ob Universalien

en auch in der realen Welt existieren, oder ob die entsprechenden Bezeichnungen stets Mengen von Objekten bezeichnen.

In den meisten Fällen ist die Interpretation von Universalien als Mengen von Objekten sinnvoll. Die Universalie 'Kunde' zum Beispiel ist durch die Menge aller Objekte, für welche die Aussage '...ist ein Kunde' wahr ist, angemessen repräsentiert. Wie wir später sehen werden, findet sich diese Behandlung von Universalien als Objektmengen auch in entsprechenden Repräsentationssprachen wieder. Es gibt jedoch auch Fälle, in denen die Universalie wie ein Objekt behandelt wird. Ein Beispiel ist der folgende Satz:

"Die Armut ist angestiegen."

Die Universalie "Armut" wird hier wie ein Objekt verwendet. Insbesondere wird eine Aussage über die Armut gemacht, der ein Wahrheitswert zugeordnet werden kann, ohne dass es möglich ist, diese auf ein konkretes Objekt abzubilden. Auch gibt es nicht wirklich eine Menge von Armutobjekten, die gemeinsam die Menge der Armut bilden könnten, dennoch lässt sich die Wahrheit der Aussage überprüfen. Es macht also Sinn, auch Universalien als mögliche Referenten im Sinne der Terminologie von Abbildung 1.1 zu betrachten, die nicht ohne weiteres als Menge von Objekten angesehen werden können. Dies bedeutet also, dass Symbole im Prinzip für konkrete Objekte, Mengen von Objekten, aber auch für abstrakte Begriffe wie eben die oben genannte Armut stehen können. In letzteren beiden Fällen sprechen wir auch von Konzepten, die in Ermangelung eines konkreten Referenzobjektes nicht immer eindeutig zu fassen sind. Eine sinnvolle Verwendung der entsprechenden Symbols ist daher nur deshalb möglich, weil, wie oben angedeutet mit dem entsprechenden Konzept eine gewisse Erwartung verbunden ist, die sich in der Regel aus der Erfahrung, bzw. der Sozialisierung des Betrachters, sowie häufig auch aus dem Kontext der Verwendung eines Begriffes ergibt.

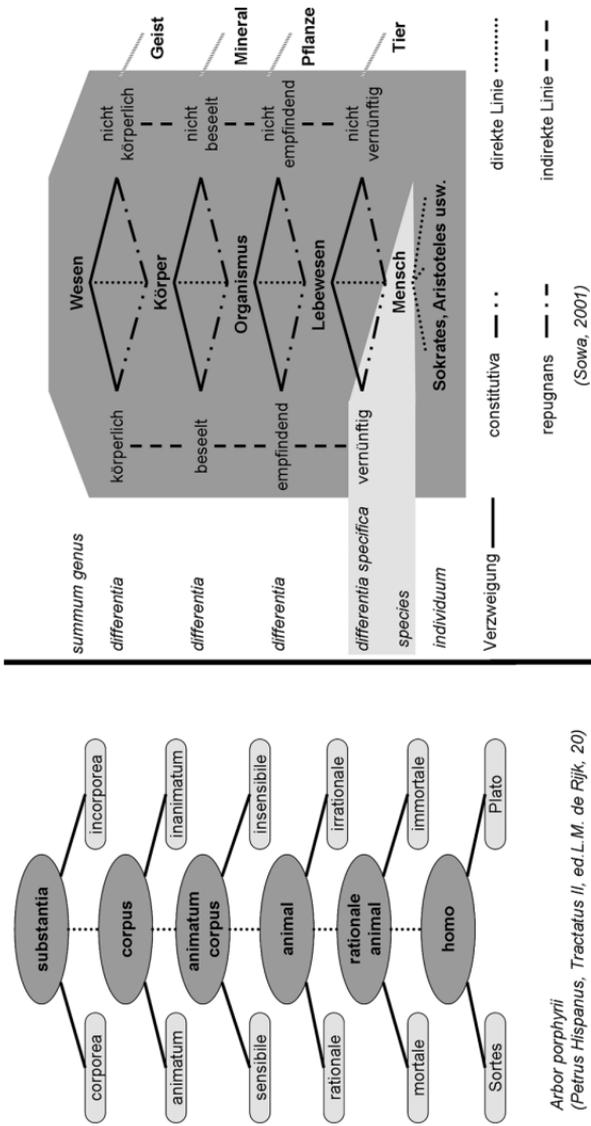
## 1.2 Objekte und Konzepte

Eine weitere Fragestellung mit der sich die Philosophie beschäftigt, ist die Frage, welche Arten von Objekten es in der realen Welt gibt. Hierbei werden Objekte in der größten allgemein denkbaren Form betrachtet. Diese werden auch als Entitäten bezeichnet und umfassen obiger Argumentation folgend, sowohl konkrete Objekte wie Personen, als auch abstrakte Objekte (Universalien) wie Organisationen, oder Konten aus unserem Beispiel. Die Betrachtungen umfassen hierbei zum einen die Frage, was allen Objekten gemeinsam ist, also im Grunde, was Existenz eigentlich bedeutet, zum anderen aber auch die Frage, was Objekten voneinander unterscheidet. Diese Unterscheidung von Objektarten ist eng verbunden mit dem dritten Element des semiotischen Dreiecks, welches die Erwartungen an die Natur eines dargestellten Objektes repräsentiert. Diese Erwartungen ergeben sich direkt aus den typischen Eigenschaften eines Objektes, welches es von anderen Arten von Objekten unterscheidet. So erwartet man von einer Person, die ein Konto besitzt, dass sie auch einen Vornamen hat. Von einer Firma als Inhaber eines Kontos, erwartet man hingegen zum Beispiel einen Eintrag ins Handelsregister. Von diesem wiederum erwartet man, dass er die entsprechende Firma eindeutig identifiziert, was man von dem Vornamen einer Person nicht erwarten würde. Die eindeutige Identifizierbarkeit über einen Eintrag ins Handelsregister ist also eine Eigenschaft, die ein Objekt von der Art Unternehmen von einem Objekt der Art Person unterscheidet. Während die Frage nach den Gemeinsamkeiten aller existierenden Objekte sich im Kontext der Informationsverarbeitung nicht stellt – es sollen ja stets Objekte abgebildet werden, von deren Existenz wir wissen oder ausgehen – ist die Frage, was Objekte und damit unsere Erwartungen an sie unterscheidet, eine wesentliche, da sie direkten Einfluss auf die Gestaltung der

Abbildung der realen Welt in eine informationstechnische Darstellung hat. Beim Entwurf einer Datenbank mit Konten und Kontoinhabern ist es, wie man leicht einsehen kann, von entscheidender Bedeutung, ob und wie ein Kontoinhaber eindeutig identifiziert werden kann. Im Falle von Unternehmen könnte dies demnach über den Eintrag ins Handelsregister erfolgen, im Falle von Privatpersonen ist eine andere Form der Identifizierung notwendig. Eine eng hiermit verbundene Frage, die auch im Rahmen der klassischen Ontologie eine Rolle spielt, ist die nach der Veränderlichkeit bestimmter Eigenschaften. Das Geburtsdatum einer Person zum Beispiel wird sich nicht mehr ändern. Es kann, eine korrekte Erfassung vorausgesetzt, für alle Zeit als gültig angesehen werden. Ganz anders verhält es sich mit dem Namen eines Unternehmens, der sich durchaus ändern kann, ohne dass dies das Unternehmen an sich in Frage stellt.

Im Gegensatz zu solchen, doch recht konkreten Überlegungen zu Objekten und ihren Eigenschaften, beschäftigt sich die Philosophie traditionell eher mit grundlegenden Unterscheidungen, wie der zwischen materiellen und immateriellen Gegenständen und versucht, sogenannte Kategoriensysteme zu entwickeln, die Objekte aufgrund ihrer Eigenschaften in unterschiedliche Kategorien einteilen. Kategoriensysteme sollen hierbei in der Philosophie in der Regel umfassend sein, d.h. Kategoriensysteme sollen die Menge aller denkbaren realen Objekte beschreiben und sich nicht auf ein bestimmtes Anwendungsgebiet, wie etwa das Electronic Banking in unserem Beispiel beschränken. Außerdem ergibt sich durch das erklärte Ziel der Unterscheidung von Objekten die Notwendigkeit, Kategorien so zu definieren, dass eine eindeutige Zuordnung von Objekten zu Kategorien besteht.

Abbildung 1.2 zeigt unterschiedliche Darstellungen des von Porphyry vorgeschlagenen Kategoriensystems für Objekte, welches auf Aristoteles zurückgeht. Obwohl dieses System inzwi-



Arbor porphyryi  
(Petrus Hispanus, Tractatus II, ed. L.M. de Rijk, 20)

Abb. 1.2 Porphyrys Kategoriensystem für 'Substance' nach Petrus Hispanus und Sowa

schen überholt ist, ist es doch nützlich, um die grundlegende Ideen eines Kategoriensystems genauer zu betrachten. Insbesondere zeigt sich an dieser Darstellung sehr schön das Prinzip von Kategoriensystemen, Objekte anhand charakteristischer Eigenschaften zu klassifizieren. Im Kategoriensystem von Porphyry werden Objekte der Kategorie 'Substanz' aufgrund bestimmter differenzierender Eigenschaften in weitere Unterkategorien aufgeteilt. Die erste dieser als 'differentiae' bezeichneten Eigenschaften ist hier die Körperlichkeit. Substanzen, die nicht körperlich sind, fallen in die Kategorie 'Geist', körperliche Substanzen werden anhand weiterer Eigenschaften klassifiziert. Die Körperlichkeit dient hierbei als eindeutiges Kriterium und dient sowohl der Zuordnung von Objekten zu einer Klasse (Ist eine Substanz nicht körperlich, fällt sie automatisch in die Kategorie Geist), als auch dem Ausschluss von Objekten von einer Kategorie (ist etwas nicht körperlich, so kann es nicht zu der Kategorie der Körper gehören). Umgekehrt erlaubt die Zugehörigkeit eines Objektes zu einer Kategorie auch Rückschlüsse auf (einige von) dessen Eigenschaften. So besitzt jedes Objekt der Kategorie "Geist" notwendigerweise die Eigenschaft, nicht körperlich zu sein. Die "differentiae" beschreiben demnach notwendige und hinreichen Bedingungen für die Zugehörigkeit zu einer bestimmten Kategorie. Ein weiteres Phänomen, welches sich gut anhand des Kategoriensystems in Abbildung 1.2 erklären lässt, ist das der Vererbung von Eigenschaften an Unterkategorien. Betrachten wir die Kategorie 'Mensch' so stellen wir fest, dass Objekte dieser Kategorie automatisch auch in den Kategorien Lebewesen, Organismus usw. enthalten sind. Diese nach unserer Intuition selbstverständliche Tatsache ergibt sich im Kategoriensystem zwangsläufig aus der oben beschriebenen Definition von Kategorien über charakteristische Eigenschaften. Wurde ein Objekt als Mensch klassifiziert, so ist dies stets Folge einer Entscheidungskette, in der die vorhande-

nen "differentiae" überprüft und das Objekt den entsprechenden Kategorien zugewiesen wird. Im Falle des Objektes 'Aristoteles' wird zunächst die Körperlichkeit überprüft. Aufgrund des positiven Ergebnisses dieser Prüfung wird Aristoteles als Körper eingestuft. Im nächsten Schritt folgt die Überprüfung der Beseeltheit, welche Aristoteles als Organismus klassifiziert, usw. Dies bedeutet aber auch, dass jedes Objekt, das schließlich der Kategorie Mensch zugeordnet wird, bereits die Tests aller übergeordneten Kategorien bestanden hat und dementsprechend ebenfalls die charakteristischen Eigenschaften dieser Kategorien erfüllt. Die Verwendung charakteristischer Eigenschaften zur Unterscheidung von Objekten ist ein Prinzip, welches sich direkt auf unser Beispiel übertragen lässt. Um entscheiden zu können, ob es notwendig ist, zwischen Personen und Organisationen zu unterscheiden sind zunächst charakteristische Eigenschaften zu bestimmen. Anschließend kann überprüft werden, ob Personen und Organisationen sich im Bezug auf diese Eigenschaften unterscheiden. In unserem Fall könnte eine maßgebliche Eigenschaft die eindeutige Identifizierbarkeit über den Namen sein.

Die Betrachtung der Prinzipien, die einem Kategoriensystem zugrunde liegen, führt uns zu zwei grundlegenden Ansätzen, Kategorien zu beschreiben: Zum einen über die Menge der Objekte, die zu der entsprechenden Kategorie gehören (extensionale Beschreibung) oder aber über die Eigenschaften, die allen Objekten der entsprechenden Kategorie gemeinsam sind (intensionale Beschreibung). Beide Arten der Beschreibung besitzen Vor- und Nachteile. Wollen wir beispielsweise das Konzept 'Mensch' beschreiben, so ist es relativ einfach, dies anhand der im System verwendeten Eigenschaften zu tun: Menschen sind körperlich, beseelt, empfindend und vernünftig. Eine Beschreibung der Kategorie ist wesentlich schwerer, da hierfür die mehr als 5 Milliarden Menschen einzeln aufgezählt werden müssten,