

Kapitel 8

Ontologische Relationen

ULF SCHWARZ UND BARRY SMITH

8.1 Was sind und was sollen formalontologische Relationen?

In Kapitel 5 wurde mit dem ontologischen Sextett eine Klassifizierung der Realität vorgestellt, die das klassische Aristotelische Viereck erweitert, indem sie Einzeldinge und Universalien in insgesamt sechs Kategorien umfasst. Zwischen den Entitäten dieser Kategorien sind ganz bestimmte Beziehungen postuliert worden, wie zum Beispiel Inhärenz, Partizipation, Instantiierung, Exemplifizierung und Charakterisierung. In diesem Kapitel soll der Frage nachgegangen werden, wie eben diese Beziehungen zwischen den Entitäten dieser Kategorien genauer charakterisiert und die sie bezeichnenden relationalen Ausdrücke eindeutig formal definiert werden können.

Diese Relationen sind echte formalontologische Relationen. Das heißt erstens, dass es *ontologische*, nicht rein logische Relationen sind. Einige Philosophen haben dahingehend argumentiert, dass sich Beziehungen dieser Art aus logischen Gründen ergeben, da doch im Anschluss an Frege und Boole Prädikate *prima facie* als das Allgemeine in der Wirklichkeit behandelt werden.¹⁴⁵ Da es aber jenseits der Aufgabe und des Vermögens der Logik ist, Aufschluss über die Beschaffenheit der Realität zu geben, muss es sich um ontologische, nicht um logische Relationen handeln.¹⁴⁶ Dass die zu untersuchenden Relationen formalontologische Beziehungen sind, heißt zweitens, dass sie *formale* Relationen sind: Sie sind nicht zusätzliche Bestandteile der Wirklichkeit, sondern sie sind dasjenige, was existierende Entitäten zu größeren Einheiten in der Realität verbindet.¹⁴⁷ Eine formalontologische Relation in diesem Sinn ist

¹⁴⁵ Vgl. z.B. Meixner 1992.

¹⁴⁶ Vgl. Smith 2005a.

¹⁴⁷ Vgl. Ceusters, Elkin und Smith 2006.

zum Beispiel die Teil-Ganzes-Beziehung, in der Arm, Bein, Kopf, etc. als Teile zu einer Einheit, einem Körper, verbunden sind. Es ist daher zum Erstellen einer Liste der wichtigen Relationen notwendig, die Wirklichkeit zu untersuchen und herauszufinden, welche Partikularien und Universalien es gibt und welche Beziehungen zwischen ihnen bestehen. Eine anschauliche Methode zum Erreichen dieses Zieles kann am besten entwickelt werden, indem wir die Ebene der *top level* Ontologie verlassen und uns auf die Ebene einer Bereichsontologie (*domain ontology*) begeben, die einen bestimmten Einzelbereich etwa der biologischen Wirklichkeit abbilden soll. Eine solche Beschränkung erlaubt es uns, am Beispiel eines klar abgesteckten Bereichs der Wissenschaft zu zeigen, welche Beziehungen zwischen Entitäten der biologischen Wirklichkeit bestehen und wie relationale Ausdrücke, die in biologischen Ontologien und Terminologie-Datenbanken (wie etwa kontrollierten Vokabularien) gebraucht werden, konsistent und eindeutig formal definiert werden können. Dazu müssen wir weitere Arten von Entitäten, wie Zeitpunkte oder räumliche Regionen (siehe dazu Kap. 6), explizit als existierend anerkennen. Diese kommen zwar im ontologischen Sextett nicht explizit vor, haben aber ihren festen Platz in der Biologie und sind damit auch für biologische Ontologien wichtig.

Die im ontologischen Sextett vorkommenden Relationen unterscheiden sich jeweils danach, ob die für sie charakteristischen Relata Universalien oder Einzeldinge sind. Wir können drei typische Fälle unterscheiden:

- \langle Universalie, Universalie \rangle : Beide Relata sind Universalien. Solcher Art sind zum Beispiel die Charakterisierungs-Relation oder die Subsumptions-Relation *is_a*, die zwischen der Universalie *Mensch* und der Universalie *Säugetier* besteht. Es gilt: *Mensch is_a Lebewesen*.
- \langle Instanz, Universalie \rangle : Das erste Relatum ist ein Einzelding, das zweite eine Universalie. Zu diesem Typ gehört zum Beispiel die Instantiierungs-Relation, die zwischen diesem bestimmten Mensch namens Peter und der Universalie *Mensch*, zwischen Peters Leben und der Universalie *Leben* besteht.
- \langle Instanz, Instanz \rangle : Beide Relata sind Einzeldinge. Dazu gehört etwa die Inhärenz oder die Partizipations-Relation, die zwischen Peters Leben und Peter besteht, oder auch – unabhängig vom ontologischen Sextett – die Teil-Ganzes-Relation auf der Ebene der Instanzen, die zwischen dieser

bestimmten Nase (Peters Nase) und diesem bestimmten Kopf (Peters Kopf) besteht.

Für das Folgende sollen zwei terminologische Konventionen vereinbart werden: Erstens werden wir, um eventuelle Mehrdeutigkeiten zu vermeiden, Ausdrücke für solche Relation *kursiv* setzen, die nur Universalien als Relata haben, und **fett** Ausdrücke für diejenigen Relationen, die unter ihren Relata mindestens ein Einzelding haben. Zweitens werden wir als Relationsausdrücke die im Bereich der angewandten Ontologie üblichen und bereits eingeführten Ausdrücke verwenden, die sich aus der englischen Sprache ableiten, wie *is_a* für die Subsumptions-Relation und **instance_of** für die Instantiierungs-Relation. Manche Relationen, wie die Partizipations-Relation und die Teil-Ganzes-Relation, kommen sowohl auf der Ebene der Instanzen als auch der Ebene der Universalien vor; wir haben es also sowohl mit **participates_in** und **part_of** zu tun als auch mit *participates_in* und *part_of*.

Die im ontologischen Sextett vorkommenden Relationen spielen auch in einer Ontologie der biologischen Wirklichkeit eine wichtige Rolle. Diejenigen Relationen, die Einzeldinge betreffen, also solche des zweiten und dritten Typs, werden im Folgenden verwendet, um Ausdrücke für Relationen auf der Ebene der Universalien zu definieren.¹⁴⁸ Diese sind es, die aus der Sicht der informationstechnischen Speicherung biomedizinischen Wissens im Mittelpunkt des Interesses stehen. Denn biologisches Wissen ist Wissen um Universalien: Nicht Fury und Black Beauty werden in Zoologiebüchern beschrieben, sondern die Art Pferd, nicht Einzeldinge mithin, sondern Universalien und deren Beziehungen zueinander.

Biologische Ontologien werden entwickelt, um als kontrollierte Vokabularien zu dienen, in denen wissenschaftliche Ergebnisse der biologischen Forschung ausgedrückt werden können. Sätze der Form „*A Relation B*“, wobei „*A*“ und „*B*“ Terme einer biologischen Ontologie sind und „*Relation*“ Platzhalter für *part_of* oder einen ähnlichen Ausdruck ist, drücken allgemeine Aussagen über die entsprechenden biologischen Universalien aus, für die diese

¹⁴⁸ Um unnötig komplizierte sprachliche Konstruktionen zu vermeiden, werden im Folgenden „Definition einer Relation“ und „Definition eines relationalen Ausdrucks“ synonym gebraucht und nur an solchen Stellen voneinander unterschieden, an denen Verwirrung droht.

Ausdrücke stehen. Welche Entitäten¹⁴⁹ als biologische Universalien gelten (und welche nicht), kann nicht so einfach beantwortet werden. Es kann jedoch zunächst darauf verwiesen werden, dass es solche Universalien sind, die durch die Verwendung der sie benennenden Ausdrücke in der biologischen Fachliteratur implizit anerkannt werden, wie *Zelle* oder *Fortpflanzung*. Feststellungen, die die Instanzen und Partikularien der biologischen Realität betreffen (zum Beispiel eine Aussage über das spezifische Gewicht dieses bestimmten Organismus in genau dieser Petrischale), sind zwar in der biologischen Forschung unverzichtbar, gehören aber nicht zu den allgemeinen Aussagen der biologischen Wissenschaften. Demnach verbinden Relationen in biologischen Ontologien Universalien miteinander. Die Ausdrücke, die in einer biologischen Ontologie Anwendung finden, benennen ausschließlich Universalien, also dasjenige, was in der Realität allgemeinen Charakter besitzt (siehe Kap. 5). Es lässt sich jedoch nicht definieren, was es heißt, dass eine Universalie in einer bestimmten Beziehung (zum Beispiel der Teil-Ganzes-Relation: *part_of*) zu einer anderen steht, ohne dabei auf die entsprechenden Instanzen der Universalien zurückzugreifen und explizit zu machen, in welchen Beziehungen diese zueinander stehen. So kann nur angegeben werden, was es bedeutet, dass *Retina part_of Auge* ist, wenn erkannt wird, dass dies eine Aussage darüber ist, dass alle Instanzen der Universalie *Retina* in einer *part_of*-Beziehung auf der Ebene der Instanzen zu einigen Instanzen der Universalie *Auge* stehen. Die Abhängigkeit der Relationen zwischen Universalien von Relationen zwischen ihren entsprechenden Instanzen bildet die Grundlage unserer Definitionen der relationalen Ausdrücke, welche die allgemeinen Termini (Namen von Universalien) miteinander verbinden. In allen zu erstellenden Definitionen muss klar zum Ausdruck kommen, dass sie nicht zufällige (kontingente), sondern gesetzmäßige Beziehungen und Verhältnisse zwischen biologischen Universalien wiedergeben.

Im Folgenden werden exemplarisch drei korrekte formale Definitionen der Klassenzugehörigkeits-Relation (*is_a*), der Teil-Ganzes-Relation (*part_of*) und der Partizipations-Relation (*has_participant*) zwischen Universalien vorgestellt

¹⁴⁹ Das Wort „Entität“ wird im Folgenden als universaler ontologischer Terminus verwendet, der alles einschließt, was in irgendeiner Weise existiert (alle Kontinuante, Prozesse, Funktionen, Strukturen, Orte, Zeiten etc., sowohl auf der Ebene der Instanzen als auch der Universalien).

und so gezeigt, wie formale Definitionen auch für andere ontologische Relationen zwischen Entitäten anderer Bereiche der Realität zu formulieren sind.¹⁵⁰

8.2 Formalontologische Relationen definieren: Chancen und Probleme

Mittels einer durchgängigen Verwendung formaler Definitionen und der damit verbundenen guten Charakterisierung formaler Relationen in Ontologien kann Interoperabilität zwischen verschiedenen Ontologien und die Unterstützung automatisierter Schlussfolgerungen gewährleistet werden. Denn die für die jeweiligen Fachbereiche entwickelten Terminologie-Systeme und die damit verbundenen Klassifizierungen müssen aufeinander abgestimmt sein, um einen gewinnbringenden Informationsaustausch zu ermöglichen. Ontologische Relationen müssen also in den verschiedenen Systemen kohärent behandelt werden, d.h. ein relationaler Ausdruck muss immer für ein und dieselbe Relation stehen, auch wenn er in verschiedenen Ontologien benutzt wird. Aber schon im Hinblick auf ontologische Grundrelationen, wie die Klassenzugehörigkeitsbeziehung und die Teil-Ganzes-Relation, ist offensichtlich, dass gleiche relationale Ausdrücke in verschiedenen Ontologien nicht notwendigerweise für äquivalente Beziehungen stehen. Diese Verwirrung geht sogar so weit, dass diese beiden Grundrelationen gar nicht voneinander unterschieden werden. So ist in der medizinischen Terminologiedatenbank UMLS (*Unified Medical Language System*) die Behauptung zulässig: *Pflanzenblätter is_a Pflanze*, im Standardsystem SNOMED eine Behauptung wie: *beide Uteri is_a Uterus*. Die hier vorgestellte Methode erlaubt es, relationale Ausdrücke zwischen allgemeinen Termini eindeutig formal zu definieren und so ihre Bedeutung präzise anzugeben.

In der gängigen Literatur zur Wissensrepräsentation¹⁵¹ werden leider oftmals Relationen zwischen Universalien als einfache Grundrelationen eingeführt, ohne weitere Spezifizierungen anzugeben. Ein solches Vorgehen lässt allerdings Wesentliches außer Acht: Es ist oftmals nicht trivial, für allgemeine

¹⁵⁰ Zum Erreichen dieses Zieles müssen alle in einer Ontologie angeführten relationalen Ausdrücke zwischen allgemeinen Termini formal definiert werden. Für eine umfassende Behandlung von biologischen Relationen vgl. Smith et al. 2005.

¹⁵¹ Vgl. Fellbaum 1998.

Termini festzustellen, ob eine bestimmte Relation zwischen den entsprechenden Universalien tatsächlich besteht. Weil es sowohl weibliche als auch männliche Menschen gibt, können wir wohl behaupten, dass *menschlicher Hoden part_of Mensch* (Teil-Ganzes-Beziehung) gilt, *Mensch has_part menschlicher Hoden* (Ganzes-Teil-Beziehung) aber nicht. Weil es Säugetiere mit Herzen gibt, die nicht menschlich sind, gilt *Mensch has_part Herz*, nicht aber *Herz part_of Mensch*. Aus entsprechenden Gründen gilt auch *Wachstum has_participant Mensch* (Participationsbeziehung) nicht. Die Zeitdimension kann ebenfalls Probleme bereiten. Es gilt wohl kaum *Erwachsener is_a Kind* (Klassenzugehörigkeits-Beziehung), weil jede Instanz der Universalie *Erwachsene* zu einem früheren Zeitpunkt einmal die Universalie *Kind* instantiiert hat.

Viele biologische Bereichsontologien enthalten leider relationale Ausdrücke für Beziehungen zwischen Universalien, die nur durch informelle Hinweise und lose formulierte Anweisungen für ihre richtige Verwendung eingeführt werden. Ohne formale Definitionen bleiben die logischen Verbindungen zwischen verschiedenen Relationen im Dunkeln und oftmals haben relationale Ausdrücke sogar innerhalb einer Ontologie verschiedene Interpretationen¹⁵², ganz davon zu schweigen, dass sie auch in verschiedenen Ontologien nicht mit einer konsistenten Bedeutung benutzt werden.¹⁵³

8.3 Arten von Relationen

Relationen in biologischen Ontologien müssen die folgenden vier Bedingungen erfüllen, damit die vorgeschlagene Methode auf sie anwendbar ist und sie in eine entsprechende Ontologie der biologischen Relationen aufgenommen werden können:

Erstens sind die Relationen, die in diesem Kapitel besprochen werden, echte ontologische Relationen. Es handelt sich dabei somit um Beziehungen, die zwischen Entitäten der Wirklichkeit bestehen und die unabhängig sind von unseren Verfahren oder Methoden, über sie Kenntnis zu erlangen. Zudem sind sie unabhängig von unseren Mitteln, sie zu repräsentieren oder über sie unser Wissen für Computer verarbeitbar zu machen. Das gilt keineswegs für

¹⁵² Vgl. Ceusters et al. 2004.

¹⁵³ Vgl. Smith, Köhler, Kumar 2004a.

alle Relationen. Beispielsweise verwendet das *Gene Ontology Annotation Project* die Relation:

„A“ *annotates* B

Sie wird verwendet, um identifizierte Gene oder Genprodukte in Verbindung zu setzen mit solchen Ausdrücken aus einer zum Indizieren von Genetik-Aufsätzen verwendeten Liste von Standard-Vokabeln (einem „kontrollierten Vokabular“), die diese Klassen von Genen oder Genprodukten benennen oder beschreiben.¹⁵⁴ Da diese Relation nicht eine Universalie zu einer anderen Universalie der Natur in Bezug setzt, sondern zu Ausdrücken in einem Vokabular, das die betroffene Wissenschaftsdisziplin selbst erzeugt hat, handelt es sich nicht um eine ontologische Beziehung (vgl. Kap. 2).

Zweitens sind hier nur solche Relationen von Interesse, die, zumindest im Prinzip, in allen biologischen Ontologien erscheinen könnten und nicht nur zwischen Entitäten sehr spezifischer Art bestehen. Ein Beispiel für eine sehr spezifische Relation wäre *A ist_Genom_von B*, die nur in einer Sequenz-Ontologie eine Bedeutung hätte.

Die Relationen müssen drittens ein hohes Maß an Allgemeingültigkeit aufweisen. Aussagen über das Bestehen von bestimmten Beziehungen zwischen Entitäten sollen entweder gesetzmäßige Aussagen sein oder sie sollen Aussagen sein, in denen eine analytische Beziehung zwischen Universalien zum Ausdruck kommt. Eine solche analytische Verbindung beruht auf der Zerlegbarkeit von Allgemeinausdrücken (Universalienamen) in Fällen wie *Hautzelle is_a Zelle* oder *Herzinfarkt has_participant Herz*.

Viertens soll es für jede Relation zwischen Universalien eine einfache formale Definition geben. Falls man sich nicht auf eine solche Definition einigen kann, so kann die Relation nicht in eine biologische Ontologie aufgenommen werden. Wir bestehen auf diesem Kriterium, weil wir Definitionen aufstellen wollen, die leicht verständlich sind und gleichzeitig auch logikbasierte Computerverfahren zum Schlussfolgern und zur Konsistenzprüfung unterstützen können.

Eine Relation, die dieses Kriterium nach heutigem Stand der biologischen Wissenschaft nicht erfüllt, ist die Beziehung der molekularen Funktion. Es ist

¹⁵⁴ Vgl. dazu: <http://www.geneontology.org/GO.annotation.shtml> (8.8.2006).

noch nicht gelungen, einfach verständliche, unumstrittene primitive Relationen anzugeben, mit deren Hilfe eine Definition der molekularen Funktion auf Universalienebene zu erstellen wäre. Die *Gene Ontology* enthält molekulare Funktionen. In ihr sind molekulare Funktionen ungefähr so etwas wie „Aktivitäten“ auf Molekülebene; als Beispiele werden unter anderem katalytische Aktivitäten oder Verknüpfungsaktivitäten angeführt. Es besteht aber wenig Einigkeit darüber, was genau die Träger dieser Aktivitäten sind und wie sich Aktivitäten von Aktionen oder Ereignissen ontologisch unterscheiden.¹⁵⁵

Wie bereits angesprochen, sollen in einer biologischen Ontologie nur solche Relationen repräsentiert werden, die die Form $\langle \text{Universalie, Universalie} \rangle$ haben. Zum Erstellen formaler Definitionen dieser Relationen wird es aber notwendig sein, auf Relationen zurückzugreifen, die von der Form $\langle \text{Instanz, Instanz} \rangle$ oder $\langle \text{Instanz, Universalie} \rangle$ sind. Denn Relationen des Typs $\langle \text{Universalie, Universalie} \rangle$ sind Beziehungen, die nur aufgrund bestimmter Beziehungen der Art $\langle \text{Instanz, Instanz} \rangle$ oder $\langle \text{Instanz, Universalie} \rangle$ zwischen ihren Instanzen oder einer Universalie und ihren Instanzen bestehen. Eine solche Charakterisierung der Beziehungen zwischen Universalien ist in der aristotelischen Interpretation des Universalienbegriffs verankert, nach der Universalien nur in ihren Instanzen existieren (vgl. Kap. 6).

8.4 Arten von Relata und Beschränkungen der Anwendung relationaler Ausdrücke

Bevor wir beginnen können, einige grundlegende relationale Ausdrücke in einer biologischen Ontologie zu definieren, müssen wir angeben, welche Ausdrücke benutzt werden sollen, um die Entitäten zu benennen, die als Relata in bestimmten Relationen zueinander stehen. Dabei müssen wir sowohl Instanzen als auch Universalien mit geeigneten Mittel anführen können. Wir werden uns dabei der Werkzeuge der Logik bedienen und dazu Variablen und Quantoren verwenden.¹⁵⁶ Es sollen unterschiedliche Sorten von Variablen als Platzhalter für Instanzen und Universalien von Kontinuanten und Prozessen sowie für Zeitpunkte verwendet werden:

¹⁵⁵ Für erste Versuche, Licht in dieses Dunkel zu bringen, vgl. Schulz und Jansen 2006.

¹⁵⁶ Vgl. z.B. Kutschera 1992.

- C, C_1, \dots stehen für Universalien von Kontinuanten;
- P, P_1, \dots stehen für Universalien von Prozessen;
- c, c_1, \dots stehen für Instanzen von Kontinuanten;
- p, p_1, \dots stehen für Instanzen von Prozessen;
- t, t_1, \dots stehen für Zeitpunkte.

Kontinuanten und Prozesse sind Kategorien, die sich gegenseitig ausschließen (siehe Kap. 5). Kontinuanten können materielle Entitäten sein, wie eine Zelle oder ein Mensch, aber auch immaterielle, wie ein Loch oder ein Kanal (vgl. Kap. 6). Immaterielle Kontinuanten haben einiges gemeinsam mit räumlichen Regionen¹⁵⁷, sind jedoch von ihnen im Hinblick darauf verschieden, dass sie immaterielle Teile von Organismen sind. Genau wie materielle Kontinuanten bewegen sie sich mit der Bewegung ihrer Träger von einer räumlichen Region zu einer anderen.

Die Biologie befasst sich hauptsächlich mit dreidimensionalen Kontinuanten, die ein oberes und ein unteres Ende, ein Innen und ein Außen sowie eine Vorder- und Rückseite haben. Prozesse hingegen haben einen Anfang, eine Mitte und ein Ende. Im Gegensatz zu Kontinuanten können Prozesse entlang der Zeitachse unterteilt werden, sodass zum Beispiel ihre Jugend und ihr Erwachsensein zeitliche Teile des Prozesses sind, der ihr Leben ist.

Es handelt sich hier also um zwei unterschiedliche, komplementäre Sichtweisen auf dieselbe Realität, wobei eine raumorientiert, die andere zeitorientiert ist. Es gibt bestimmte logische und ontologische Beziehungen zwischen diesen Sichtweisen, die wir in unserer Behandlung der Relationen durch die Bezugnahme auf räumliche Regionen und auf Zeitpunkte explizit machen. Allerdings folgen daraus je nach Relation auch eindeutige Beschränkungen der Relata. So ist es beispielsweise falsch, eine relationale Aussage der Form P *is_a* C zu bilden, weil die Klassenzugehörigkeitsrelation keine Entitäten aus inkompatiblen Kategorien verbinden kann. Im Gegensatz dazu erfordert die *has_participant* Relation, dass die erstgenannte Entität eine Prozess-Universalie ist und die zweitgenannte eine Kontinuanten-Universalie, sodass diese Relation so etwas wie eine Brücke zwischen den beiden Sichtweisen auf die Realität darstellt.

¹⁵⁷ Vgl. Casati und Varzi 1994.

Die Verortung der Relata im ontologischen Sextett legt zugleich Beschränkungen fest, welche Arten von Relationen zwischen welchen Relata bestehen können. Eine Instantiierungs-Relation kann nur zwischen einem Einzelding und einer Universalie bestehen, eine Partizipations-Relation nur zwischen einem Kontinuant und einem Prozess etc. Daher werden wir im Folgenden eine Unterscheidung in zwei Arten von Beziehungen auf der Ebene der Instanzen treffen: Es gibt Relationen zwischen Kontinuanten, deren Repräsentation einen Bezug auf Zeitpunkte enthalten muss, und es gibt Relationen zwischen Prozessen, bei deren Repräsentation dies nicht erforderlich ist. Da Prozesse sich in der Zeit ereignen, tragen sie den zeitlichen Aspekt bereits in sich.

8.5 Primitive Relationen auf der Ebene der Instanzen

Um nun die Klassenzugehörigkeits-, Teil-Ganzes- und Partizipations-Relation auf der Ebene der Universalien definieren zu können, müssen zunächst diejenigen Relationen angegeben werden, die wir als nicht weiter definierbar und damit als primitiv ansehen wollen. Ansonsten droht entweder ein unendlicher Regress, wenn wir immer neue Vokabeln zur Definition der bisher vorhandenen einführen, oder eine zirkuläre Struktur, wenn wir einen Ausdruck mit Hilfe von Ausdrücken definieren, in deren Definition er bereits vorkommt. Die primitiven Relationen sollten evident, selbsterklärend und neutral in Bezug zu den Wirklichkeitsbereichen (*domains*) der Einzeldisziplinen sein, sodass sie nicht nur innerhalb einzelner Bereiche der Biologie Anwendung finden können. Die Entitäten in spezifischen biologischen Ontologien stehen in Beziehungen zueinander, die auch in anderen Bereichen der Biologie oder der Wirklichkeit zwischen Entitäten bestehen können. Bis auf die Instantiierungsbeziehung (**instance_of**), die zwischen einer Instanz und einer Universalie besteht, sind alle primitiven Relationen Beziehungen zwischen Instanzen von Universalien, die wir dann benutzen können, um die Beziehungen zwischen den Universalien selbst zu definieren.

Wir wählen die folgenden Relationen als primitiv:

- *c* **instance_of** *C* at *t* – eine primitive Relation zwischen einer Kontinuanten-Instanz und einer Universalie, die sie zu einem bestimmten Zeitpunkt instantiiert. Diese Relation entspricht der Instantiierungsbeziehung im ontologischen Sextett, die zwischen einem substantiellen Einzelding und einer substantiellen Universalie besteht.

- p **instance_of** P – eine primitive Relation zwischen einer Prozess-Instanz und einer Universalie, die sie instantiiert, unabhängig von der Zeit. Diese Relation entspricht der Instantiierungsbeziehung im ontologischen Sextett, die zwischen einem individuellen Prozess und einer Prozess-Universalie besteht.
- c **part_of** c_1 **at** t – eine primitive Teil-Ganzes-Relation zwischen zwei Kontinuanten-Instanzen und einem Zeitpunkt, zu dem die eine Kontinuanten-Instanz ein Teil der anderen ist.
- p **part_of** p_1 – eine primitive Teil-Ganzes-Relation, die, unabhängig von der Zeit, zwischen Prozess-Instanzen besteht (wobei einer ein Teilprozess des anderen ist).
- p **has_participant** c **at** t – eine primitive Relation der Partizipation zwischen einem Prozess, einem Kontinuanten und einem Zeitpunkt. Dies ist die inverse Relation zur Partizipationsbeziehung im ontologischen Sextett, die zu einer bestimmten Zeit zwischen einem substantiellen Einzelding und einem individuellen Prozess besteht.

Für einen menschlichen Leser sind diese Relationen relativ leicht verständlich. Um sie allerdings in Computeranwendungen benutzen zu können, muss die Bedeutung der relationalen Ausdrücke formal durch Axiome charakterisiert werden. Die Arbeit an diesen Axiomen ist noch nicht abgeschlossen, es lassen sich aber schon einige wichtige nennen.

Für die Relation **instance_of** zum Beispiel gelten folgende Axiome:

- Die Relation **instance_of** bezieht sich immer auf eine Instanz und eine Universalie, in dieser Reihenfolge.
- Keine Entität kann zugleich Instanz und Universalie sein.

Für die Relation **part_of** gelten folgende Axiome:¹⁵⁸

- Die Relation **part_of** ist irreflexiv: Keine Entität ist ein Teil von sich selbst.
- Sie ist asymmetrisch: Wenn x in der **part_of**-Relation zu y steht, dann steht y nicht in der **part_of**-Relation zu x .

¹⁵⁸ Vgl. ausführlicher Simons 1987.

- Sie ist transitiv: Wenn x zu y und y zu z jeweils in der **part_of**-Relation stehen, dann steht x in der **part_of**-Relation zu z .
- Es gilt ein Summierungsprinzip, das die Existenz mereologischer Summen oder Ganzheiten garantiert.
- Es gilt ein Differenzierungsprinzip: Wenn x in der **part_of**-Relation zu y steht, dann existiert ein weiterer Teil z von y , der mit x keine gemeinsamen Teile hat.

Bei der Anwendung dieser Axiome auf Teil-Ganzes-Beziehungen zwischen Kontinuanten müssen diese so modifiziert werden, dass sie zeitliche Indizes enthalten.

8.6 Formale Definitionen der Relationen auf der Ebene der Universalien

Jetzt steht das Instrumentarium zur Verfügung, mit dem wir uns an die Definition relationaler Ausdrücke machen können, die Relationen zwischen Universalien bezeichnen.

8.6.1 Die Definition der Klassenzugehörigkeits-Relation

Die *is_a*-Relation kann mit der Beziehung der Klassenzugehörigkeit identifiziert werden, die aus der mathematischen Mengentheorie bekannt ist. Die Relation **instance_of** kann nach dieser Auslegung auch als Entsprechung zur mengentheoretischen Beziehung der Klassenmitgliedschaft angesehen werden. Dann wäre eine Definition von $A \text{ is_a } B$ in folgender Art und Weise denkbar:

$A \text{ is_a } B =_{\text{def}}$ Für alle x gilt: wenn x **instance_of** A , dann x **instance_of** B .

Leider kann diese Auslegung höchstens eine notwendige Bedingung für die Wahrheit von $A \text{ is_a } B$ liefern. Zwei Argumente sprechen dagegen, dass es sich hierbei auch um eine hinreichende Bedingung handelt:

(1) Diese Auslegung erlaubt Fälle, in denen die Klassenzugehörigkeit nur kontingent ist, wie zum Beispiel:

Zellkern in 10 ml-Reagenzglas is_a Zellkern.

(2) In dieser Auslegung findet der zeitliche Aspekt keine Berücksichtigung. Das führt dazu, dass falsche Resultate erzeugt werden, wenn man sie auf Kontinuanten-Universalien anwendet, wie zum Beispiel:

Erwachsener is_a Kind.

Diese Beziehung besteht nicht, obwohl jede Instanz von *Erwachsener* zu einem Zeitpunkt auch Instanz von *Kind* war.

Wir können das Argument (1) entschärfen, indem wir darauf verweisen, dass wir nur solche Beziehungen als *is_a*-Beziehungen erlauben wollen (vgl. Kap. 8.6), die in Form von Aussagen ausgedrückt werden, die Wahrheiten der biologischen Wissenschaft sind. Das sind Aussagen, die Namen von echten biologischen Universalien (wie „Enzym“ oder „Zelle“) enthalten und keine kommerziellen oder indexikalischen Termini (wie „Zellkern in diesem Reagenzglas“).

Argument (2) kann entkräftet werden, indem der zeitliche Aspekt in *is_a*-Relationen zwischen Kontinuanten-Universalien berücksichtigt wird. Kontinuanten können im Laufe ihrer Existenz, im Gegensatz zu Prozessen, verschiedene Universalien instantiieren und dabei ihre Identität bewahren. Wir müssen daher zwei Arten von *is_a*-Relationen unterscheiden: Die *is_a*-Relation zwischen Kontinuanten enthält einen zeitlichen Index, die *is_a*-Relation zwischen Prozessen ist zeitunabhängig. (Kap. 9 wird dafür argumentieren, dass es eine noch weit größere Vielfalt von *is_a*-Relationen gibt, die an dieser Stelle aber noch nicht relevant ist.) Wir definieren also:

- i. $C \text{ is_a } C_1 =_{\text{def}}$ Für alle c, t , wenn c **instance_of** C at t ,
dann c **instance_of** C_1 at t .
- ii. $P \text{ is_a } P_1 =_{\text{def}}$ Für alle p , wenn p **instance_of** P ,
dann p **instance_of** P_1 .

8.6.2 Die Definition der Teil-Ganzes-Relation

Auch bei der Teil-Ganzes-Beziehung auf der Ebene der Universalien müssen zwei Arten von *part_of*-Relationen unterschieden werden, je nachdem, ob sie zwischen Kontinuanten oder Prozessen besteht. Für Kontinuanten gilt: $C \text{ part_of } C_1$ genau dann, wenn jede Instanz von C zu jedem Zeitpunkt in der *part_of*-Relation (auf der Ebene der Instanzen) zu einigen Instanzen von C_1 zu diesem Zeitpunkt steht. So zum Beispiel: *Zellkern part_of Zelle*. Wir definieren also:

- i. $C \text{ part_of } C_1 =_{\text{def}}$ Für alle c, t , wenn c **instance_of** C at t ,
dann gibt es ein c_1 , sodass gilt:
 c_1 **instance_of** C_1 at t und c **part_of** c_1 at t .

$C \text{ part_of } C_1$ besagt, dass Instanzen von C , wann immer sie existieren, als Teile von Instanzen von C_1 existieren. Analog dazu gilt für Prozesse: $P \text{ part_of } P_1$ genau dann, wenn auf der Ebene der Instanzen jede Instanz von P in der Relation **part_of** zu mindestens einer Instanz von P_1 steht. So zum Beispiel: *Jugend part_of Leben*. Wir definieren also:

- ii. $P \text{ part_of } P_1 =_{\text{def}}$ Für alle p , wenn p **instance_of** P ,
dann gibt es ein p_1 , sodass gilt:
 p_1 **instance_of** P_1 und p **part_of** p_1 .

$P \text{ part_of } P_1$ besagt, dass Instanzen von P immer als Teile von Instanzen von P_1 existieren.

Die hier verwendeten Definitionen (bzw. ihr jeweiliges Definiens) haben eine gemeinsame logische Struktur: Sie bestehen beide aus einer Allaussage und einer Existenzaussage, die in einer Implikationsbeziehung zueinander stehen. Eine Allaussage ist eine Aussage, die einen Allquantor enthält („für alle x gilt, dass ...“) enthält, und eine Existenzaussage ist eine Aussage, die den Existenzquantor enthält („es gibt mindestens ein y , für das gilt ...“) enthält. Dass sie in einer Implikationsbeziehung zueinander stehen heißt, dass sie durch eine „wenn ..., dann ...“-Struktur miteinander verbunden sind. Diese logische Form nennt man „Alle-Einige-Struktur“ (*all-some-structure*). Sie ist verantwortlich für bestimmte logische Eigenschaften, in denen Aussagen über Teil-Ganzes-Beziehungen zwischen Universalien zueinander stehen. So kann aus *menschliche Gebärmutter part_of Mensch* nicht gefolgert werden, dass *Mensch* als Teil *menschliche Gebärmutter* hat. Denn während alle Instanzen von *menschliche Gebärmutter* zu einem bestimmten Zeitpunkt Teil von einigen Instanzen von *Mensch* sind, müssen zu diesem Zeitpunkt nicht alle Instanzen von *Mensch* auch Instanzen von *menschliche Gebärmutter* als Teile haben.

8.6.3 Die Definition der Partizipations-Relation

Die primitive **has_participant** Relation auf der Ebene der Instanzen besteht zwischen einem Prozess, einem Kontinuanten und einem Zeitpunkt, zu dem der Kontinuant auf irgendeine Art und Weise an dem Prozesse beteiligt ist. Diese Beziehung besteht zum Beispiel, wenn an diesem bestimmten Prozess eines Zelltransportes diese bestimmte Zelle zu diesem bestimmten Zeitpunkt beteiligt ist. Zur Definition der **has_participant** Beziehung auf der Ebene der Universalien verfahren wir analog zu oben. Wir definieren also:

$P \text{ has_participant } C =_{\text{def}}$ Für alle p gilt: wenn p **instance_of** P ,
dann gibt es ein c und ein t , sodass gilt:
 c **instance_of** C , und p **has_participant** c at t .

Auch hier ist anzumerken, dass $P \text{ has_participant } C$ lediglich besagt, dass Instanzen von P Instanzen von C als Träger oder Beteiligte benötigen. Aufgrund der Alle-Einige-Struktur der Definition von has_participant kann daraus aber nicht gefolgert werden, dass Instanzen von C immer an Prozessen einer bestimmten Art beteiligt seien. So folgt aus *menschliches Fortpflanzungsverhalten has_participant Mensch* keineswegs, dass alle Menschen an menschlichem Fortpflanzungsverhalten teilnehmen.

8.7 Logik der Relationen

Die inverse Relation („Umkehrrelation“) R^{-1} zu einer zweistelligen Relation R ist definiert als die Relation, die zwischen einem Paar von Relata genau dann besteht, wenn die ursprüngliche Relation R zwischen denselben Relata in umgekehrter Reihenfolge besteht. Zu den von uns diskutierten primitiven Relationen auf der Ebene der Instanzen können leicht inverse Gegenstücke definiert werden. Auch die Definition der inversen Relation zu is_a ist trivial:

$A \text{ has_subclass } B =_{\text{def}} B \text{ is_a } A$.

Allerdings bringt die has_subclass -Relation keinen Gewinn hinsichtlich der Ausdrucksfähigkeit der Ontologie mit sich: Jede Information, die sich mit Hilfe von has_subclass ausdrücken lässt, konnte zuvor schon mit is_a ausgedrückt werden. Ähnlich uninformativ sind die inversen Relationen zu den übrigen Relationen auf der Ebene der Universalien, deren Definition eine Alle-Einige-Struktur hat. Deren inverse Relationen haben zudem selbst keine Alle-Einige-Struktur mehr: Obwohl auf der Ebene der Instanzen „ x **is_part_of** y “ genau dann wahr ist, wenn die mittels der inversen Relation formulierte Aussage „ y **has_part** x “ wahr ist, gilt analoges auf der Ebene der Universalien nicht: So ist beispielsweise die relationale Aussage *menschliche Hoden part_of Mensch* wahr, denn jede Instanz von *menschlicher Hoden* ist Teil einer Instanz der Universalie *Mensch*, aber es gibt keine entsprechende Relation mit Alle-Einige-Struktur, die jede Instanz von *Mensch* mit mindestens einer Instanz von *menschlicher Hoden* als Teile von ihnen verbindet. Das ist nicht etwa als Defizit unserer Definitionen der Relationen anzusehen, sondern spiegelt lediglich die Beschaffenheit der Realität wider: Nicht alle Instan-

zen der Universalie *Mensch* haben eine Instanz der Universalie *menschliche Hoden* als Teil.

Trotzdem ist eine *has_part*-Relation auf Universalien-Ebene sehr erwünscht, denn mit ihr könnten Sachverhalte wie *Mensch has_part Herz* oder *Zelle has_part Zellkern* ausgedrückt werden. Diese Relationen auf Universalien-Ebene, die zwar keine inversen Relationen sind, aber mit Hilfe derselben Relationen auf der Instanzen-Ebene in einer Alle-Einige-Struktur definiert werden, wie ihre Gegenstücke, wollen wir „reziproke Relationen“ nennen. Für die zwei *part_of*-Relationen ergeben sich zum Beispiel die Definitionen der entsprechenden reziproken Relationen so:

$C \text{ has_part } C_1 \quad =_{\text{def}} \quad \text{Für alle } c \text{ und } t \text{ gilt: wenn } c \text{ instance_of } C \text{ at } t, \\ \text{dann gibt es ein } c_1, \text{ sodass gilt:} \\ c_1 \text{ instance_of } C_1 \text{ at } t \text{ und } c_1 \text{ part_of } c \text{ at } t.$

$P \text{ has_part } P_1 \quad =_{\text{def}} \quad \text{Für alle } p \text{ gilt: wenn } p \text{ instance_of } P, \\ \text{dann gibt es ein } p_1, \text{ sodass gilt:} \\ p_1 \text{ instance_of } P_1 \text{ und } p_1 \text{ part_of } p$

Da es sich dabei nicht um inverse sondern nur um reziproke Relationen handelt, kann aus *A part_of B* nicht auf *B has_part A* geschlossen werden. Genauso wenig kann von *A has_part B* auf *B part_of A* gefolgert werden. Aus der wahren Aussage *menschliche Hoden part_of Mensch* folgt nicht, dass *Mensch has_part menschliche Hoden* gilt, und aus *Zellwand part_of Zelle* folgt nicht, dass *Zelle has_part Zellwand* gilt. Dass es sich hier nicht um Paare von inversen Relationen handelt, sondern um Paare von reziproken Relationen, bringt zwar diese Einschränkung der Folgerungsmöglichkeiten mit sich, aber zugleich auch einen Zugewinn an Ausdruckskraft. Denn, wie hinsichtlich von *is_a* und *has_subclass* bereits erwähnt, lässt sich alles, was mit einer inversen Relation ausgedrückt werden kann, auch mit der ursprünglichen Relation zum Ausdruck bringen. Eine reziproke Relation bringt hingegen ein Plus an Ausdruckskraft mit sich, das es ermöglicht, Beziehungen zwischen Universalien auszudrücken, die mit der ursprünglichen Relation nicht formulierbar waren.

Die Eigenschaften der definierten Relationen auf der Ebene der Universalie im Hinblick auf Transitivität, Reflexivität und Symmetrie der Relationen sind in der Tabelle in Abbildung 8.1 zusammengestellt.

Abb. 8.1: Logische Eigenschaften einiger formaler Relationen

Relation	transitiv	symmetrisch	reflexiv	antisymmetrisch
<i>is_a</i>	+	-	+	+
<i>part_of</i>	+	-	+	+
<i>has_participant</i>	-	-	-	-

8.8 Ausblick

Durch die klaren formalen Definitionen der relationalen Ausdrücke gelingt es, ihre Bedeutung klar anzugeben. So wird eindeutig, für welche Beziehung zum Beispiel ein relationaler Ausdruck wie *is_a* steht, und zwar in einer Form, die nicht nur Beschränkungen, was die möglichen Relata betrifft, sondern auch logische Eigenschaften der Beziehungen zwischen Universalien innerhalb einer Ontologie explizit macht. Die Methode, die wir zum Erstellen der Definitionen verwenden, ist für alle Spezialdisziplinen der Biologie und damit für alle biologischen Bereichsontologien geeignet und auch auf andere Gebiete der Wissenschaft übertragbar. Unsere Definitionen sind so angelegt, dass sie eine einheitliche Behandlung der entsprechenden relationalen Ausdrücke in allen biologischen Ontologien erlauben. Auf diese Weise wollen wir einen Beitrag dazu leisten, die Interoperabilität unter verschiedenen Ontologien zu fördern, auch wenn diese zu völlig verschiedenen Zwecken und von unterschiedlichen wissenschaftlichen Disziplinen entwickelt werden. Unsere Methode zum Erstellen der Definitionen gewährleistet auch, dass die Anwendung automatisierter Schlussfolgerungsmechanismen zur Informationsgewinnung aus Ontologien mit definierten Relationen möglich wird. Dazu muss allerdings noch einige Arbeit geleistet werden, was die Axiomatisierung der primitiven Relationen betrifft.

Letztlich ist es die Aufgabe der Einzelwissenschaften, festzulegen, welche Relationen in ihrem Bereich als primitiv anzusehen sind und wie aus ihnen Definitionen der relationalen Ausdrücke zwischen allgemeinen Termini ihres Fachbereiches zu erstellen sind. Für biomedizinische Ontologien sind über die hier diskutierten Relationen hinaus bereits formale Definitionen für zehn relationale Ausdrücke vorgeschlagen worden, die in der Biologie und Medizin häufig Anwendung finden, darunter

- die räumlichen Relationen *located_in*, *contained_in* und *adjacent_to* zwischen Kontinuanten,
- die Relationen *derives_from* und *transformation_of* zwischen Kontinuanten und
- die Relationen *preceded_by* für Prozesse und *has_agent* für Prozesse und Kontinuanten.

An der Definition weiterer formaler Definitionen wird intensiv gearbeitet. Zu den Desiderata der biomedizinischen Ontologie gehört es, Definitionen von Relationen wie die Folgenden formal zu definieren:¹⁵⁹

- *A lacks B* (*A fehlt B*)
- *A connected_to B* (*A verbunden_mit B*)
- *A realized_by B* (*A realisiert_durch B*)
- *A function_of B* (*A Funktion_von B*)
- *A has_enabler B* (*A ermöglicht_durch B*)
- *A has_result B* (*A hat_Ergebnis B*)

Obwohl je nach dem darzustellenden Gegenstandsbereich einige andere primitive Relationen zwischen Instanzen von Bedeutung sein werden, so werden doch die primitiven Grundrelationen zwischen Entitäten des ontologischen Sextetts immer eine zentrale Rolle spielen, weil sie die grundlegenden Beziehungen zwischen Entitäten der Realität in ihrer allgemeinsten Form darstellen. Auch wenn wir uns hier mit einigen Relation und Anwendungen beschäftigt haben, die für die Biologie und Medizin wichtig sind, kann die von uns angewandte Methode der Definition auch zum Erstellen von formalen Definitionen der relationalen Ausdrücke für andere Gegenstandsbereich der Wissenschaften verwendet werden.

¹⁵⁹ Für Vorschläge zur Formalisierung weiterer Relationen vgl. Smith und Grenon 2004 und Schulz und Hahn 2007.