

Basisontologie zur semantischen Annotation von Nomina – Manual*

Sarah Metzger, Sebastian Bücking, Frauke Buscher, Johanna Herdtfelder, Julia Lukassek,
Anna Prysłowska, Sarah Zobel, Claudia Maienborn

Inhalt

1. Einleitung	1
2. Die A1- Ontologie	2
2.1 Vorbilder	2
2.1.1 SUMO	2
2.1.2 DOLCE	5
2.2 Die A1-Ontologie – ein einführender Überblick	8
2.2.1 Kategorien	9
2.2.2 Merkmale (Features)	11
3. Anwendung der Annotationsschemata	12
3.1 Allgemeine Richtlinien der Annotation	12
3.1.1 Gegenstand der Annotationen – Nomen oder NP?	13
3.1.2 Die Berücksichtigung des Bedeutungszusammenhangs im Kontext	14
3.2 Der Entscheidungsbaum	15
3.2.1 Die Konzeption des Entscheidungsbaums	16
3.2.2 Die Verwendung des Entscheidungsbaums	16
3.3 Der Umgang mit uneindeutigen Fällen	18
4. Literatur	19

1. Einleitung

- (1) a. Die Professorin begann die Vorlesung.
b. ?Die Professorin begann den Pullover.

Während in (1a) ein Ereignisargument vorliegt, das eine reibungslose Interpretation erlaubt, ist der Pullover in (1b) ein Artefakt, das in ein Ereignisargument uminterpretiert werden muss. Daraus resultiert, dass der semantische Typ der Argumente relevant für die Interpretation eines Satzes ist und ein sehr nützliches Instrumentarium zur Erfassung der Kombinatorik und der Bedeutung sprachlicher Ausdrücke darstellt (Maienborn 2017). Insbesondere Vertreter der

* Die hier vorgestellte Basisontologie und die dazugehörigen sprachlichen Diagnostiken wurden im Rahmen des DFG-Projektes „Kombinatorische Bedeutungsanpassung an der Semantik-/ Pragmatik-Schnittstelle“ des SFB 833 „Bedeutungskonstitution“ der Eberhard Karls Universität Tübingen konzipiert.

konzeptuellen Semantik greifen auf eine feinkörnige semantische Typenhierarchie zurück; für einen Überblick siehe Maienborn (2017). So werden beispielsweise sowohl in Pustejovskys (1991, 1995) Generativem Lexikon als auch in Ashers (2011) Typenkompositionslogik semantische Typen zur Repräsentation von Bedeutung verwendet. Um mit diesen semantischen Typen jedoch Bedeutung adäquat modellieren zu können, ist es wichtig, semantische Typen auf einer möglichst objektiven Grundlage zu definieren. Des Weiteren erfordert die Bedeutungsmodellierung anhand semantischer Typen klare sprachliche Diagnostiken, die eine objektive und eindeutige Zuordnung von sprachlichen Ausdrücken zu semantischen Typen erlauben.

Die vorliegende Ontologie soll genau dies leisten. Während die Ontologie selbst die ontologischen Zusammenhänge verschiedener semantischer Typen in einer Hierarchie darstellt, dienen die in Form eines Rasters dargestellten Diagnostiken dazu, konkretes sprachliches Material in diese Ontologie einzuordnen. Ziel ist es dabei, die in Korpusdaten enthaltenen Argumente (genauer: Nominalphrasen) auf einer theoretisch fundierten Grundlage zu klassifizieren und diese Kategorien als möglichst objektive Bezeichnungen in Korpusstudien verwenden zu können.

Der Aufbau des Handbuchs ist wie folgt: Abschnitt 2 liefert zunächst einen Überblick über die Ontologie (inklusive vorangehender Ansätze, die zur Konzeption herangezogen wurden). Abschnitt 3 illustriert die Anwendung der Annotationsschemata und enthält Anmerkungen zur Annotation, wie zum Beispiel die Einbeziehung von sprachlichem Kontext. Dabei werden auch schwierige Fälle, deren Klassifikation nicht eindeutig erscheint, diskutiert.

2. Die im A1-Projekt entwickelte Basisontologie

2.1 Vorbilder

Unsere Ontologie wurde in Anlehnung an bereits bestehende Basisontologien entwickelt: die *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE) (Masolo et al. 2003), und die *Suggested Upper Merged Ontology* (SUMO) (Niles & Pease 2002; Pease 2006). Die Entitäten, die in diesen Ontologien angenommen werden, dienten als Ausgangspunkt für unsere Überlegungen. Wer sich ausschließlich für die A1-Ontologie interessiert, kann diesen Abschnitt überspringen und auf Seite 8 bei Abschnitt 2.2 weiterlesen.

2.1.1 SUMO

Abbildung 1 zeigt die *Suggested Upper Merged Ontology*, kurz: SUMO. Die SUMO-Ontologie wurde entwickelt, um den Anforderungen moderner digitaler Standards gerecht zu werden: E-Commerce und Softwareintegration erfordern spezifische und präzise Begriffe, die dem Aussagenpotenzial einer natürlichen Sprache ähneln und gleichzeitig von Computern

interpretiert werden können (Niles & Pease 2002). Jeder Versuch von Vertretern aus der Linguistik, der Kognitionswissenschaft, der Philosophie, der Künstlichen Intelligenz sowie der Bibliothekswissenschaften allein scheiterte aus mangelndem Konsens mit anderen Disziplinen oder aus mangelnder Fähigkeit zur Formalisierung (2002). SUMO ist eine interdisziplinär entwickelte Ontologie, die als Grundlage für fachspezifische Domänenontologien dienen soll; siehe Pease et al. (2006) speziell für die Word-Net-basierte Anwendung in der Linguistik. Die aus wenigen Verzweigungen bestehende Taxonomie wurde auf die wichtigsten Kategorien komprimiert.

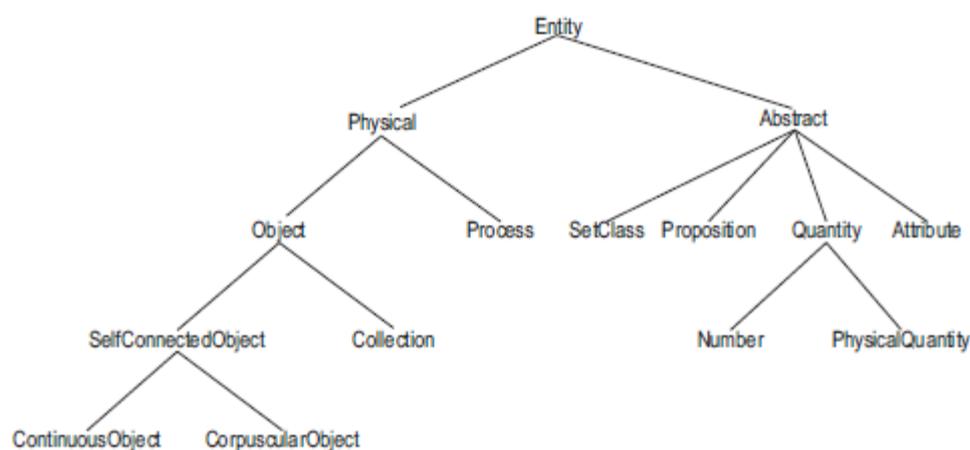


Abb. 1: The Suggested Upper Merged Ontology (SUMO) (Oberle et al. 2006: 8)

Im Folgenden soll der grundlegende Aufbau nach Niles & Pease (2002) nochmals skizziert werden: Die Hauptunterscheidung wird zwischen physikalischen und abstrakten Entitäten gemacht. Physikalische Entitäten können in Raum und Zeit positioniert werden und lassen sich in Objekte und Prozesse unterteilen. Während Erstere zu jeder Zeit als Ganzes vorhanden sind, gilt dies für Letztere nicht. Prozesse entwickeln sich somit in der Zeit. Unter den Objekten wird zwischen dem selbst-zusammengefügten (*self-connected*) Objekt und der Kollektion unterschieden. Kollektionen bestehen aus nicht-zusammengehörigen Teilen, deren Relation untereinander auf Mitgliedschaft beruht; z.B. *Schafherde*, *Werkzeugkiste*. Selbst-zusammengefügte Objekte sind Objekte, deren Teile direkt oder indirekt miteinander verbunden sind. Sie lassen sich in korpuskulare Objekte (z.B. *Schaf*, *Frau*, *Kind*) und beständige Objekte unterteilen. Ein beständiges Objekt liegt vor, wenn die Teile dieselben Eigenschaften aufweisen wie das Ganze, z.B. Substanzen wie *Lehm* und *Wasser*, geographische Gebiete und Oberflächen. Anders als physikalische Entitäten können abstrakte Entitäten nicht in Raum und Zeit positioniert werden. Darunter werden Mengen (*Sets*), Propositionen, Mengenangaben und Attribute zusammengefasst. Die mengentheoretische Menge (*Set*) beinhaltet Klassen und Relationen. Eine Klasse ist

eine Menge, deren Mitglieder eine (oder mehrere) bestimmte Eigenschaften auszeichnet (z.B. *alle Frauen, Männer, Künstler, ...*). Eine Relation ist eine Klasse von geordneten Tupeln (z.B. *alle Mütter, Söhne, Hausbesitzer, ...*). Die Proposition bezieht sich auf den semantischen Gehalt, wobei nicht spezifiziert wird, ob es sich dabei um einen Satz oder ein ganzes Buch handelt. Attribute umfassen sämtliche Eigenschaften, Merkmale, etc., die nicht als Objekt reifiziert werden (z.B. *männlich* oder *weiblich*). Die Mengenangaben werden in Anzahl und physikalische Mengen unterteilt. Bei der Anzahl handelt es sich um eine von einem Maßsystem unabhängige Zählung (z.B. *Unmengen, eine Vielzahl, ...*) und die Mengenangabe stellt eine Zahl in Kombination mit einer Maßangabe dar (z.B. *drei Meter, 30°C, 40 kg*).

Folgende Punkte wurden aus der Ontologie übernommen: Aus der Gruppe der Abstrakta wurden Propositionen sowie Mengen und die Unterteilung in Anzahl und physikalische Mengen übernommen. Darüber hinaus wurden die Objekte und Prozesse, wenn auch mit einigen Modifikationen, übernommen. Es wurde beispielsweise in unserer Ontologie keine klare Trennlinie zwischen physikalischen und abstrakten Entitäten gezogen, da es auch abstrakte Objekte gibt. Prozesse wurden in unserer Ontologie feinkörniger ausdifferenziert. Von der Festlegung von Attributen als Abstrakta wurde abgesehen. Auf eine im Voraus festgelegte Unterscheidung von Kollektionen und selbst-zusammengefügten Objekten verzichtet. Dasselbe gilt für die Unterscheidung von korpuskularem und beständigem Objekt.

Auf spezifische Problematiken der SUMO-Ontologie soll im Folgenden eingegangen werden: Objekte werden in SUMO unter die physikalischen Entitäten subsumiert. Dies ignoriert die Tatsache, dass es auch abstrakte Objekte wie beispielsweise *Ideen* gibt. Darüber hinaus werden Eigenschaften bzw. Merkmale (*attributes*) als Abstrakta klassifiziert. Wir gehen jedoch mit Moltmann (2013: 300) davon aus, dass es auch trägergebundene, konkrete Eigenschaftsmanifestationen, die sogenannten Tropen, gibt. Deshalb wurde auf eine Festlegung der Eigenschaften als Abstrakta verzichtet. Darüber hinaus ist in der SUMO-Taxonomie zwar die Kategorie *Prozess* (*process*), jedoch nicht die Kategorie *Zustand* (*state*) vorhanden. Die Unterteilung in korpuskulares Objekt und Kollektion berücksichtigt nicht, dass Kollektiva selbst oftmals aus einzelnen (korpuskularen) Objekten aufgebaut sind. Ein von Niles & Pease (2002: 8) angeführtes Beispiel ist *Schafherde*. Wir erfassen Kollektivität daher mittels semantischer Merkmale; siehe Abschnitt 2.2.2 für eine ausführliche Begründung. Eine Unterscheidung zwischen belebten und unbelebten Entitäten erlaubt die SUMO-Kategorisierung nicht. Doch da sich die Belebtheit auf die Verbinterpretation auswirken kann, ist diese Unterscheidung in einer linguistischen/ semantischen Ontologie elementar; für ein Beispiel siehe die *default*-Interpretationen bei stativ-/ eventiv-ambigen Verben (Lukassek 2012). Semantische Restriktionen, die sich aus der

Belebtheit des Subjekts ergeben, könnten mittels der Unterscheidung von korpuskularen Objekten (*Schaf, Hammer*) und Kollektionen (*Schafherde, Werkzeugkiste*) folglich nicht erfasst werden:

(2) a. Das Schaf/ die Schafherde frisst Gras.

b. *Der Hammer/ die Werkzeugkiste frisst Gras.

Um das Merkmal der Kollektivität, das auf verschiedene Arten von Objekten zutreffen kann, dennoch zu erfassen, wurde in unserer Ontologie das Merkmal *+/-collective* eingeführt. Um Substanzbegriffe (wie *Sand, Teig, ...*) abbilden zu können, wurde in unserer Ontologie das Merkmal *+/- countable* angenommen.

2.1.2 DOLCE

DOLCE ist die *Descriptive Ontology of Linguistic and Cognitive Engineering*. Sie wurde entwickelt, um eine effiziente Kommunikation innerhalb des Semantic Web zu gewährleisten (Masolo et al. 2003). Ziel ist es dabei, Begriffe in einer nicht-ambigen und maschinell verarbeitbaren Sprache zugänglich zu machen sowie bestehende Ontologien vergleichbar zu machen (Masolo et al. 2003). DOLCE soll dabei auch als Ausgangspunkt für neue Ontologien dienen (Masolo et al. 2003). Die Ontologie verfügt im Vergleich zu SUMO über eine etwas ausdifferenziertere und daher umfangreichere Taxonomie. Diese wird in Abbildung 2 dargestellt:

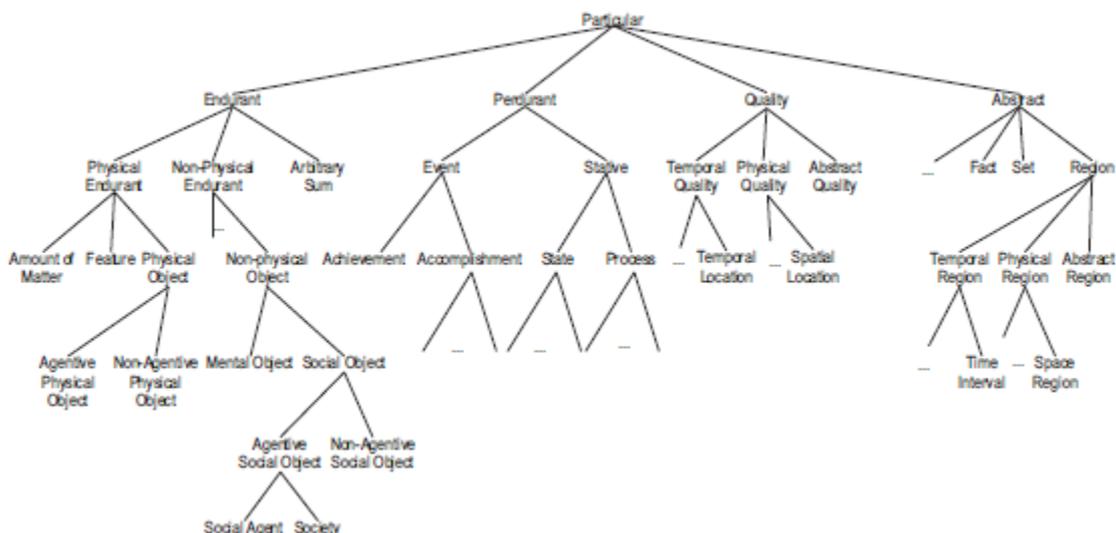


Abb. 2: *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE) (Oberle et al. 2006)

Im Folgenden soll die DOLCE-Ontologie nach Masolo et al. (2003) erläutert werden: Die Oberkategorien umfassen *Enduranten, Perduranten, Eigenschaften* und *Abstrakta*. Die Enduranten sind zu jedem Zeitpunkt ihrer Existenz vollständig vorhandene Entitäten, die sich mit der Zeit verändern können, jedoch keine zeitlichen Bestandteile haben. Sie lassen sich in physikalische und nicht-physikalische Enduranten sowie arbiträre Summen unterteilen. Ein

physikalischer Endurant weist im Gegensatz zum physikalischen Enduranten direkte räumliche, messbare und wahrnehmbare Eigenschaften auf. Bei arbiträren Summen handelt es sich um eine willkürliche Verbindung zweier Nomina, die verschiedene semantische Typen konstituieren; z.B. *mein linker Fuß und mein Auto*. Die physikalischen Enduranten werden in Materie (*amount of matter*), sogenannte *features* und physikalische Objekte unterteilt. Materielle Begriffe bilden oft kein Ganzes, weshalb sie oft Massenomen wie *Gold, Holz, Sand* oder *Eisen* bezeichnen. Features hängen stets von einem Träger ab und können nicht unabhängig existieren; z.B. *Loch, Oberfläche*, etc. Unter den physikalischen Objekten gibt es agentive und nicht-agentive Vertreter. Während die agentiven Objekte (z.B. sämtliche menschlichen Individuen) Intentionen, Wünsche und Glaubensüberzeugungen haben, gilt dies für Letztere (z.B. *Hammer, Haus, Körper*) nicht. Nicht-physikalische Enduranten umfassen sämtliche nicht-physikalischen Objekte. Diese können in mentale und soziale Objekte unterteilt werden. Mentale Objekte (z.B. eine *Empfindung*) sind nicht von einer Gemeinschaft abhängig, während soziale Objekte dies sind. Letztere lassen sich in agentive und nicht-agentive nicht-physikalische Objekte unterteilen. Bei nicht-agentiven sozialen Objekten handelt es sich um Begriffe wie *Gesetze*. Bei den agentiven physikalischen Objekten wird zwischen den sozialen Aktanten (z.B. der *US-Präsident*) und Gesellschaften (z.B. der *SWR*) unterschieden. Die Perduranten passieren in der Zeit und haben somit zeitliche Bestandteile. Es gibt eventive Perduranten wie Achievements (z.B. *Abfahrt*) und Accomplishments (z.B. *Konferenz*) sowie stative Perduranten, die Prozesse (z.B. *das Rennen*) und Zustände (*das Sitzen*) subsumieren. Die Eigenschaften werden in DOLCE als Tropen analysiert. Sie sind grundlegende Entitäten, die anderen Entitäten inhärent und von diesen abhängig sind bzw. eine direkte oder indirekte räumliche Ausprägung haben. Sie können wahrgenommen und gemessen werden. In DOLCE wird zwischen physikalischen, nicht-physikalischen und zeitlichen Eigenschaften unterschieden. Physikalische Eigenschaften (z.B. *Gewicht*) sind physikalischen Objekten inhärent, nicht-physikalische Eigenschaften (z.B. *Vermögenswert*) beschreiben nicht-physikalische Objekte und zeitliche Eigenschaften (z.B. *die Dauer des Zweiten Weltkriegs*) beziehen sich auf Perduranten. Unter die Klasse der Abstrakta, die keine räumlichen oder zeitlichen Eigenschaften aufweisen, fallen Fakten, Mengen und Regionen. Erstere haben jedoch eine rein bezeichnende Funktion, während Letztere bestimmte Werte im Eigenschaftsraum (d.h. konkrete Ausprägungen einer Eigenschaft) darstellen; z.B. *die Zeitachse, 22.06.2002, eine Sekunde, 80 kg, ein Ort, ein Bereich auf dem Farbspektrum, der Wert eines Euros*).

Wir haben die DOLCE-Taxonomie in Grundzügen übernommen. Als besonders gelungen sollen folgende Aspekte hervorgehoben werden: In DOLCE fallen abstrakte Entitäten wie *Idee*

unter nicht-physikalische Objekte und gehören somit derselben Oberklasse an wie physikalische Objekte, nämlich den Enduranten. Damit kann erfasst werden, dass beide Objektarten in denselben sprachlichen Kontexten auftreten können:

(3) a. Der Präsident/ die Kanzlerin/ der Chef hat eine Entscheidung getroffen.

b. Das Parlament/ die Schule/ die Firma/ die Partei hat eine Entscheidung getroffen.

Zum anderen werden bei DOLCE Eigenschaften als Tropen und nicht als Abstrakta definiert (Masolo et al. 2003: 11). Dies trägt der Tatsache Rechnung, dass Tropen trägergebunden und daher konkret wahrnehmbar sein können, während dies für Abstrakta nicht gilt (Moltmann 2013: 300).

Um jedoch die Komplexität der Taxonomie zu reduzieren, wurden auf für unsere Zwecke irrelevante Verzweigungen verzichtet. So wurden die Enduranten auf die Kategorie der Objekte reduziert, die wiederum in physikalische und nicht-physikalische Objekte unterteilt werden. Auf die Kategorie *Arbitrary Sum* wurde verzichtet, da diese arbiträre Kombination nichts über die semantischen Typen der Argumente aussagt. Für unsere Zwecke verzichtbar ist die Kategorie *Amount of Matter*, da diese Kategorie nichts darüber aussagt, von welchem Typ die bezeichnete Entität ist. Relevant ist hingegen die Unterscheidung zwischen natürlichem Objekt und Artefakt, da sie Unterschiede in der Verbsemantik erfassen kann (siehe (4)). Deshalb sollte die Unterscheidung von Artefakten und natürlichen Objekten auf jeden Fall in die Ontologie integriert und gegenüber des Substanzbegriffs priorisiert werden.

(4) a. John baked the potato. (Pustejovsky 1991: 415)

b. John baked a cake. (Pustejovsky 1991: 415)

Während ein natürliches Objekt eine Zustandswechsellesart nach Pustejovsky (1991: 415) auslöst, kommt es in Kombination mit einem natürlichen Objekt zu einer Kurationslesart. Um das Merkmal „Masse“ dennoch zu erfassen, wurde in unserer Ontologie das Merkmal *+/-countable* eingeführt.

Auf die Kategorie *Feature*, unter die laut Masolo et al. (2003: 10) Entitäten wie Löcher, Grenzen, etc. fallen, wurde verzichtet. Da diese stets trägerabhängig sind – eine Grenze muss stets die Grenze von etwas sein, während sich ein Loch stets in einer Art von Materie manifestieren muss – könnte man diese gegebenenfalls als Tropen oder physikalische Objekte analysieren. Eine endgültige Festlegung ist jedoch vorerst nicht möglich.

Die nicht-physikalischen Objekte werden in mentale und soziale Objekte unterteilt. Letztere werden als *agentive* und *nicht-agentive soziale Objekte* subkategorisiert. Da DOLCE (außer die Abhängigkeit von Aktanten) keine verlässlichen Kriterien zur Unterscheidung von mentalen und nicht-agentiven sozialen Objekten anbietet, haben wir uns dafür entschieden, nur die

agentiven sozialen Objekte (als Institution) zu klassifizieren. Um dennoch nicht-physikalische Objekte wie *Idee* annotieren zu können, die keine Institutionen sind, nehmen wir *Inhaltsobjekte* an. Darunter fallen zum einen als Unterkategorie die mentalen Objekte und die Info-Objekte. Während der konzeptuelle Gehalt der mentalen Objekte mit einem *dass*-Satz oder einem *zu*-Infinitiv spezifiziert werden kann, sind Info-Objekte Entitäten, die man je nach Repräsentationsformat entweder malen oder schreiben bzw. lesen kann:

- (5) a. Die Vorstellung/ der Gedanke/ die Idee dass...
- b. Das Glück/ die Möglichkeit/ die Pflicht zu x-en
- (6) a. Maria las/ schrieb das Buch/ den Text/ die Anzeige/ ...
- b. Paul malte ein Bild/ ein Porträt/ ein Plakat/ ...

Die *Perduranten* werden in DOLCE in stative und eventive Entitäten unterteilt. Als stative Entitäten werden auch *Prozesse* kategorisiert, obwohl diese Entitäten eigentlich dynamisch sind (Maienborn 2015: 33). Die Zustände werden nicht weiter differenziert. Wir nehmen für unsere Ontologie in Anlehnung an Maienborn (2003, 2005, 2007) *D-Zustände* und *K-Zustände* an. Hinsichtlich der Eigenschaften haben wir ebenfalls ein paar Änderungen vorgenommen: So wird auf eine genaue Ausdifferenzierung von zeitlichen, physikalischen und nicht-physikalischen Eigenschaften verzichtet.

Was die Abstrakta betrifft, wurden einige Änderungen hinsichtlich der *Regions* vorgenommen. Während die *temporal Regions* als zeitliche Entitäten erhalten bleiben, wurde für *abstract Regions* wie *der Wert eines Euros* (Masolo et al. 2003: 10) die SUMO-Kategorie *Amount* mit ihren Subklassen *Anzahl* und *Maßeinheit* gewählt. Darüber hinaus wurden in Anlehnung an DOLCE *Facts* in die Kategorie der Abstrakta aufgenommen.

Unsere Taxonomie ähnelt in den Grundzügen sehr stark der DOLCE-Taxonomie. Ein entscheidender Unterschied besteht jedoch darin, dass die Komplexität der Taxonomie auf der einen Seite so weit wie möglich reduziert wurde, um diese möglichst einfach und überschaubar zu halten. Auf der anderen Seite wurde die Taxonomie um wichtige Konzepte wie *Info-Objekte*, *Artefakte*, *natürliche Objekte*, etc. erweitert.

2.2 Die A1-Ontologie – ein einführender Überblick

Die hier entwickelte Ontologie entstand auf der Grundlage von SUMO und insbesondere DOLCE; für eine komplette Darstellung siehe Anlage 1. Anders als diese Ontologien unterscheiden wir jedoch systematisch zwischen Kategorien und Merkmalen (*features*). Kategorien bezeichnen dabei die ontologischen Entitäten, die im Rahmen unserer Basisontologie angenommen werden. Merkmale bezeichnen semantische Informationen, die auf verschiedene

Kategorien zutreffen können und eventuell eine Rolle bezüglich der Kombinatorik einiger Verben spielen können.

2.2.1 Kategorien

Abbildung 3 zeigt die angenommenen Oberkategorien:

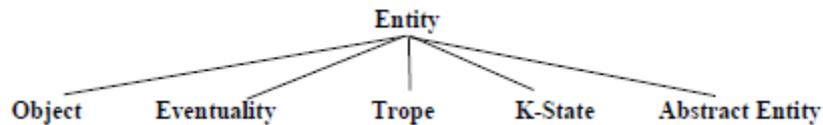


Abb. 3: Oberkategorien der Basisontologie des Projektes A1 des SFB 833 (Eberhard Karls Universität Tübingen) Prinzipiell wurden die Oberkategorien nach DOLCE übernommen. Ein zentraler Unterschied ist, dass der Grad der Abstraktheit von links nach rechts zunimmt. Dies ist wichtig für die Kategorisierung sprachlicher Belege im Entscheidungsbaum; siehe Abschnitt 3.2.

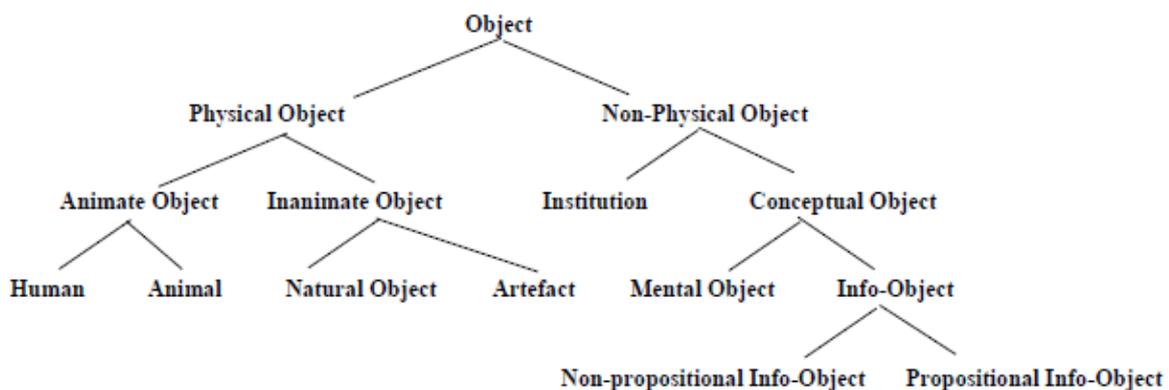


Abb. 4: Objekte der Basisontologie des Projektes A1 des SFB 833 (Eberhard Karls Universität Tübingen) Die Objekte werden – je nach Präsenz bzw. Absenz lokaler Eigenschaften (Masolo et al. 2003: 16) –in physikalische und nicht physikalische Objekte unterteilt. Die physikalischen Objekte können entweder belebt (Menschen und Tiere; z.B. *Schüler, Polizist, Mädchen, Tiger*) oder unbelebt (Artefakte und natürliche Objekte, z.B. *Hammer, Tisch, Sand*) sein. Bei den nicht-physikalischen Objekten unterscheiden wir zunächst Institutionen, also reglementierte Einrichtungen, Vereinigungen oder Gremien (z.B. *die Uni Tübingen, die SPD, der SWR*), und Inhaltsobjekte (*conceptual object*), die in irgendeiner Hinsicht einen Inhalt spezifizieren. Letztere werden in mentale Objekte (z.B. *Gedanke, Idee, Traum, Glück, Pflicht*) und propositionale Info-Objekte (z.B. *Text, Gedicht, Buch, Broschüre*) oder nicht-propositionale Info-Objekte (z.B. *Bild, Plakat, Porträt*) unterteilt. Sie unterscheiden sich dadurch, dass man bei den mentalen Objekten einen mittels *zu*-Infinitiv oder *dass*-Satz spezifizierten Inhalt explizieren kann; siehe (5), hier wiederholt als (7).

- (7) a. Die Vorstellung/ der Gedanke/ die Idee dass...
- b. Das Glück/ die Möglichkeit/ die Pflicht zu x-en

Die Eventualitäten, Entitäten mit einer raumzeitlichen Ausdehnung (Maienborn 2015: 28), zeichnen sich dadurch aus, dass sie sich in der Zeit abspielen (Masolo et al. 2003: 10).

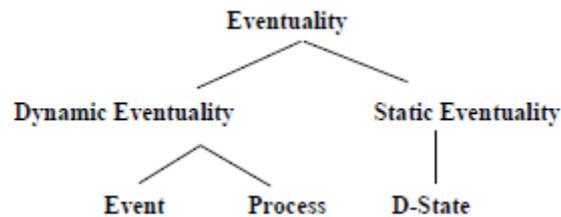


Abb. 5: Eventualitäten der Basisontologie des Projektes A1 des SFB 833 (Eberhard Karls Universität Tübingen) Bei den Eventualitäten wird zunächst zwischen statischen Eventualitäten (D-Zuständen, wie z.B. *das Warten, das Stehen*) und dynamischen Entitäten unterschieden. D-Zustände sind statische Eventualitäten, die über ein Davidson'sches Ereignisargument e_s verfügen (Maienborn 2015: 61ff.). Die dynamischen Eventualitäten können in Ereignisse und Prozesse unterteilt werden. Sie gehen auf die Situationstypen nach Vendler (1967) zurück. Prozesse (z.B. *das Rennen*) sind atelisch, während Ereignisse (wie z.B. *Ankunft, Unfall*) telisch sind und somit einen Kulminationspunkt aufweisen (Maienborn 2015: 33).

Eine weitere Kategorie sind die Tropen. Dabei handelt es sich um Eigenschaften, die eine Trägerentität aufweist und die von dieser abhängig sind (Masolo et al. 2003: 11). Sie sind typischerweise wahrnehmbar und zum Teil auch messbar (ebd.). Moltmann (2013: 300) charakterisiert sie daher als „partikularisierte Eigenschaftsmanifestationen in einem Individuum“.

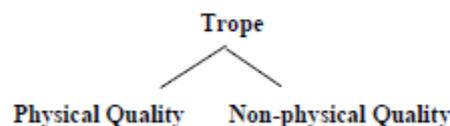


Abb. 6: Tropen der Basisontologie des Projektes A1 des SFB 833 (Eberhard Karls Universität Tübingen) Tropen können sich aus einer physikalischen (z.B. *Röte, Größe, Kälte*) oder einer nicht-physikalischen Eigenschaft (z.B. *Preis*) konstituieren.

Rechts von den Eigenschaften sind die sogenannten K-Zustände (z.B. *das Wissen der Antwort, das Kennen der Wahrheit*) angesiedelt. Diese verfügen, anders als die D-Zustände, nicht über ein Davidson'sches Ereignisargument und sind somit abstrakter (Maienborn 2017).

Darauf folgen die abstrakten Entitäten, mental konstruierte Entitäten, die kognitiven Operationen dienen, jedoch nicht unabhängig von diesen existieren (Asher 1993).

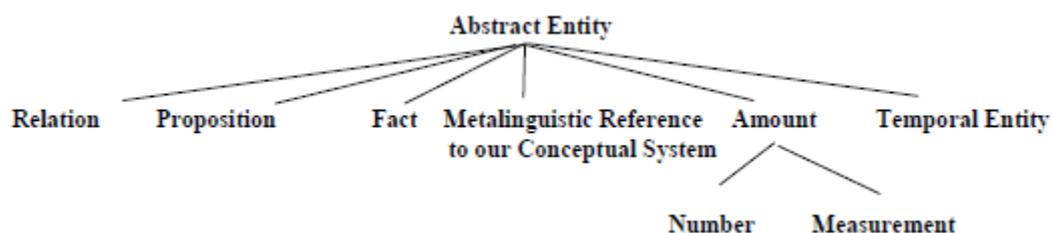


Abb. 7: Abstrakta der Basisontologie des Projektes A1 des SFB 833 (Eberhard Karls Universität Tübingen)
 Diese werden in *Relationen*, *Propositionen*, *Fakten*, *metalinguistische Bezüge zu unserem Begriffssystem*, *zeitliche Entitäten* und *Mengen* eingeteilt. Letztere lassen sich zudem in *Anzahl* (z.B. *ein Dutzend*) und *Maßeinheiten* (z.B. *Meter*, etc.) unterteilen. Unter *Relationen* werden sämtliche Begriffe zusammengefasst, die einen Gegenstand auf einen anderen abbilden (z.B. *Funktion*, *Idempotenz*). *Metasprachliche Begriffe zu unserem Begriffssystem* sind in der Regel Buchstaben, Zahlen sowie andere Zeichen (z.B. @, 5, ?, β). Es ist davon auszugehen, dass diese Kategorien in einem Korpus eher selten zugewiesen werden müssen. Aus Gründen der Vollständigkeit werden sie jedoch trotzdem angeführt.

2.2.2 Merkmale (Features)

Eine weitere wichtige Komponente der vorliegenden Ontologie sind die sogenannten *Features*. Diese können je nach Bedarf kategorienunabhängig eingesetzt werden. Wir nehmen dabei die folgenden Features mit binären Ausprägungen an: *+/-countable*, *+/-collective*, *+/-iterative*, *+/-manifested in an event*, *+/-habitual*, *+/-intentional*. Zur besseren Verständlichkeit seien an dieser Stelle noch ein paar Beispiele angeführt, die die Nützlichkeit der Features illustrieren. Das Merkmal *+/-countable* kann für verschiedene Arten von Entitäten wichtig sein, wie die folgenden Belege verdeutlichen:

- (8) a. (*zwei/ *ein paar) Sand vs. (zwei/ ein paar) Bäume (natürliches Objekt)
- b. (*zwei/ *ein paar) Beton vs. (zwei/ ein paar) Betonplatten (Artefakt)

Wie in (8) ersichtlich wird, können verschiedene Arten von Objekten sowohl zählbar als auch nicht zählbar sein. Dies erfordert ein flexibel einsetzbares Merkmal. Ähnliches gilt auch für das Merkmal *+/-collective*:

- (9) a. Schaf vs. Schafherde (belebtes physikalisches Objekt)
- b. Hammer vs. Werkzeugkiste (Artefakt)

Das Feature *+/-iterative* ist dann nützlich, wenn Situationen iteriert werden. Dies kann jedoch für Ereignisse, Prozesse und Zustände gleichermaßen gelten und nicht gelten, weshalb es nicht sinnvoll ist, Iterativität als ontologische Eigenschaft anzunehmen.

- (10) a. das zweimalige Essen von Kartoffelsuppe (Accomplishment)
- b. das dreimalige Gewinnen beim Schachspielen (Achievement)
- c. das mehrmalige Lachen (Prozess)
- d. das wiederholte Kranksein (Zustand)

Dieses Merkmal kann insbesondere dann nützlich sein, wenn ein punktuelles Verb in einem durativen sprachlichen Kontext auftritt:

- (11) das jahrelange Erreichen sämtlicher Ziele

Während *Erreichen* eigentlich ein Ereignis denotiert, deutet *jahrelang* darauf hin, dass dieses Ereignis über eine längere Zeitspanne hinweg iteriert wurde. Analog dazu kann auch das Merkmal *+/-habitual* sinnvoll sein, wenn etwas gewohnheitsmäßig erfolgt:

- (12)
- a. das häufige Essen von Kartoffelsuppe (Accomplishment)
 - b. das ständige Gewinnen beim Schachspielen (Achievement)
 - c. das tägliche Lachen (Prozess)
 - d. das ständige Kranksein (Zustand)

Darüber hinaus gibt es beispielsweise Eigenschaften, die sich in einem Ereignis manifestieren, wie zum Beispiel *Rücksichtslosigkeit*. Jemand kann also nur rücksichtslos sein, indem er rücksichtslose Handlungen vollzieht. Bei Bedarf kann das Feature *+/-manifested in an event* angenommen werden, wenn dieses für die Fragestellung relevant sein könnte. Das Merkmal *+/-intentional* kann Aufschluss darüber geben, ob eine Entität intentional handeln kann und somit agensfähig ist oder nicht. Auch dieses Merkmal kann auf Entitäten verschiedener Kategorien zutreffen, z.B. Artefakte oder belebte Objekte:

- (13) Das Kind/der Roboter hat absichtlich das Beet zertrampelt.

Betont werden muss, dass Features, wenn sie der Fragestellung dienlich sind, hinzugefügt werden können, es jedoch nicht müssen. Ob man Features für bestimmte Kategorien annotieren möchte oder nicht, kann man entweder vor dem Hintergrund der Fragestellung vor der Annotation festlegen oder man kann die Features auch später zusätzlich annotieren, wenn sich herausstellt, dass es sinnvoll sein könnte.

3. Anwendung der Annotationsschemata

Der folgende Abschnitt beschäftigt sich mit Fragen, die sich häufig zu Beginn eines Annotationsvorgangs ergeben. Dabei sollen Richtlinien darüber, inwiefern der Kontext bzw. Kotext miteinbezogen werden soll und Hinweise, wie mit uneindeutigen Fällen umgegangen werden soll, dargelegt werden. Anschließend wird die Anwendung des Entscheidungsbaums sowie den darin enthaltenen Diagnostiken exemplarisch skizziert.

3.1 Allgemeine Richtlinien der Annotation

Was soll annotiert werden – Nomen oder NP? Wie flexibel darf ein Annotator bei komplexen Nomina verfahren? Inwiefern muss sprachlicher Kontext und der daraus resultierende Bedeutungszusammenhang berücksichtigt werden? Der folgende Abschnitt gibt Aufschluss über diese Fragen.

3.1.1 Gegenstand der Annotationen – Nomen oder NP?

Zunächst einmal gilt es zu klären, welche Einheiten überhaupt annotiert werden sollen. Grundsätzlich ist es wichtig, stets die komplette NP im Blick zu behalten, da diese das Argument, dessen semantischer Typ unser primäres Erkenntnisinteresse ist, konstituiert. Es sollen daher nicht nur einzelne Nomina in Isolation betrachtet werden. Dennoch muss man sich bei komplexen Nomina bewusst machen, dass das Kopfnomen natürlich primär von Interesse ist, da dieses der Phrase den semantischen Typ zuweist. Zudem kann es sinnvoll sein, attributive Erweiterungen bei den Diagnostiken zu ignorieren. Bei Eigennamen soll zudem stets das zugrunde liegende Konzept annotiert werden, insofern dies kontextuell verfügbar ist. Diese Grundsätze sollen im Folgenden in Form von Handlungsanweisungen näher erläutert werden:

Regel 1: Annotiere stets die komplette NP.

Für unsere Fragestellungen interessant ist der semantische Typ der Argumente. Deshalb ist es sinnvoll, das Argument als Ganzes zu typisieren. Das bedeutet, dass die komplette NP berücksichtigt werden sollte. Im folgenden Beispiel wenden wir die Tests auf *die Bewohner* an, da es sich dabei um das Subjektargument von *öffnen* handelt.

- (14) Li flieht aus der Stadt, und als die Steppenreiter wieder vor Peking erscheinen, öffnen die Bewohner die Tore und begrüßen sie als Befreier.

(*Die Zeit*, 09.02.1996, Nr. 07)

Regel 2: Wenn auf den ersten Blick unklar ist, wie wir die komplexe NP annotieren sollen, dann fokussiere bei der Annotation das Kopfnomen.

- (15) Besonders die wachsende Zahl der Alleinlebenden verursacht mehr Wohn- und Heizkosten.

(*Berliner Zeitung*, 11.10.2005)

Das Subjektnomen des Verbs *verursachen*, *die wachsende Zahl der Alleinlebenden*, ist ein komplexes Nomen. In diesem Fall fokussieren wir bei der Annotation das Kopfnomen *Zahl*. Es wäre demnach falsch, in diesem Fall den semantischen Typ von *Alleinlebenden* zu bestimmen. Stattdessen konzentrieren wir uns primär auf *die wachsende Zahl*.

Regel 3: Wenn nötig, ignoriere Attribute.

Bei Bedarf kann in (15) nicht nur das Genitivattribut *der Alleinlebenden*, sondern auch das Adjektivattribut *wachsende* ignoriert werden. Dies kann z.B. dann sinnvoll sein, wenn ein Test auf dem Einsetzen eines zusätzlichen Adjektivattributs basiert.

Regel 4: Wenn nötig, konzentriere Dich im Fall von Komposita auf das Kopfnomen.

Fällt beispielsweise die Zuordnung eines Kompositums zu einer bestimmten Kategorie schwer, so ist es zulässig, probeweise(!) das Kopflexem zu annotieren. Dies wäre z.B. im folgenden Beleg hinsichtlich des Subjektnomens *Comeback-Chance* relevant:

- (16) Auf den erstarkten FC Augsburg wartet die ganz große Comeback-Chance.

(*Die Zeit*, 07.03.2013 (online))

Der Grund hierfür besteht darin, dass eine Comeback-Chance eine Art von Chance ist und die beiden Begriffe denselben semantischen Typen aufweisen sollten. Ist man sich bei der Annotation eines Kompositums unsicher, so ist es akzeptabel, den Entscheidungsbaum mittels einer NP zu durchlaufen, die lediglich das Kopfnomen statt des kompletten Kompositums enthält. In diesem Fall könnte man die Annotation probeweise(!) mit der NP *die (ganz große) Chance* durchführen.

Regel 5: Bei Eigennamen annotiere die zugrunde liegende Entität.

Es ist nicht automatisch davon auszugehen, dass sich hinter Eigennamen menschliche Individuen verstecken. Im folgenden Beispiel deuten darauf bereits die Anführungsstriche hin:

- (17) "Xaver" verändert bei Werder Bremen den vorgesehenen Tagesablauf.

(*Die Zeit*, 05.12.2013 (online))

(Vorkontext: Nach Angaben des Deutschen Wetterdienstes könnte es einer der schwersten Stürme seit Jahrzehnten in Deutschland werden.)

Berücksichtigt man den Vor- oder Nachkontext, so stellt man fest, dass es sich bei Xaver um „eine[n] der schwersten Stürme seit Jahrzehnten“ handelt. Folglich soll „Xaver“ auch als ein solcher annotiert werden.

Regel 6: Liefert der Vor- oder Nachkontext keinen Hinweis auf die Identität des Eigennamens, so schließe den Beleg aus.

3.1.2 Die Berücksichtigung des Bedeutungszusammenhangs im Kontext

Es sollte stets die NP im konkreten Bedeutungszusammenhang berücksichtigt werden. Für dot-Typen gilt: Wenn der Kontext eine Desambiguierung hinsichtlich des semantischen Typen erlaubt, sollte diese vorgenommen werden. Wenn nicht, ist die Annahme eines Dot-Typs möglich. Bei generischen Nomina ist anzunehmen, dass sie ein abstraktes Konzept repräsentieren. Daher ist eine Annotation als physikalisches Objekt nicht sinnvoll.

Regel 7: Wenn der Kontext eine eindeutige Typenzuordnung erlaubt, so weise einen eindeutigen Typ zu. Vermeide die Annotation von Dot-Typen zugunsten einer eindeutigen, kontextuell präferierten Typenzuordnung.

Bei mehrdeutigen Nomina kann der Kontext einen bestimmten Aspekt eines Nomens selektieren. Besonders evident ist dies bei sogenannten Dot-Typen (Pustejovsky 1995). Ein Dot-Typ wird einer Entität zugewiesen, die zwei Aspekte in sich vereint (Maienborn 2017). Ein Beispiel hierfür wäre z.B. *Roman* als Info-Objekt (18a) und als physikalisches Objekt (18b).

- (18) a. Dieser Roman vermittelt die Bekanntschaft mit einem anderen Gorenstein.

(Berliner Zeitung, 06.07.1995)

b. Der Roman ist bunt und wiegt ca. 100 g.

Regel 8: Erlaubt der Kontext keine eindeutige Desambiguierung zwischen mehreren Bedeutungsaspekten, annotiere die NP als Dot-Typ.

Regel 9: Wenn eine generische Lesart vorliegt, annotiere kein physikalisches Objekt, sondern ein nicht-propositionales Info-Objekt.

Intuitiv würde man *der Mensch* als belebtes physikalisches Objekt, bzw. als Mensch annotieren. Dies läuft der Interpretation des folgenden Satzes jedoch zuwider:

- (19) Hinsichtlich der aktiven Anpassungsreaktion besteht der Unterschied also darin, daß der magische Mensch inbildgestaltend nur seine Vorstellungsumwelt, der höher entwickelte Mensch aber außenweltgestaltend seine außerleibliche Dingumwelt verändert.

(Scheidt, Walter: *Kulturbiologie*, Jena: Fischer 1930, S. 77)

Der bestimmte Artikel löst im Satzkontext eine generische Interpretation des magischen Menschen als eine Art von Mensch aus. Der Mensch kann daher nicht als körperlich wahrnehmbares Individuum verstanden werden. Folglich muss der Annotator hier eine abstraktere Entität bzw. ein mentales Konzept annehmen. Eine mögliche Kategorie hierfür wäre ein nicht-propositionales Info-Objekt.

3.2 Der Entscheidungsbaum

Bei dem Entscheidungsbaum handelt es sich um eine hierarchisch angeordnete Tabelle, deren Zweck es ist, die Annotation von Nomina zu erleichtern. Der Entscheidungsbaum basiert auf Diagnostiken, die auf der Grundlage ontologischer Eigenschaften der einzelnen Kategorien entwickelt wurden. Die Genauigkeit wurde empirisch im Rahmen unseres Projekts überprüft.^{1,2} Im Folgenden soll die Konzeption sowie die Überprüfung des Entscheidungsbaums dargelegt werden. Im Anschluss daran soll die Verwendung des Baumes zunächst erläutert und exemplarisch durchexerziert werden.

¹ Überprüft wurden die Diagnostiken aufgrund der Annotation von 100 aus dem DWDS-Korpus randomisiert zusammengestellten Belegen. Die Übereinstimmung der Annotationen wurde mittels des sogenannten Fleiss-Kappa ermittelt. Der ermittelte Wert betrug je nach Durchgang und annotierter Spalte (Subjekt vs. Objekt) zwischen 0.505 und 0.718. Leider gibt es keine akzeptierte Interpretation des Kappa-Werts (Sarah Zobel, persönliche Mitteilung), weshalb man keine eindeutigen Aussagen über die Qualität der Diagnostiken machen kann. Dennoch liefern Landis & Koch (1977) eine Interpretationsrichtlinie; demnach handelt es sich in dem Bereich von 0.41 bis 0.6 um eine moderate, in dem Bereich von 0.61-0.8 sogar um eine substanzielle Übereinstimmung (1977: 165). Trotzdem deutet dies darauf hin, dass die ontologische Klasse des Referenten eine sehr anspruchsvolle Annotationskategorie ist. Man sollte daher nicht erwarten, dass Annotationen uneingeschränkt reliable Ergebnisse liefern werden. Insbesondere unerfahrene Annotatoren müssen gut instruiert werden.

² Besonderer Dank gilt Sarah Zobel, die die Annotation ausgewertet, aufbereitet und interpretiert hat.

3.2.1 Die Konzeption des Entscheidungsbaums

Der Entscheidungsbaum bildet verschiedene Fragen ab, die bei der Bestimmung der semantischen Klasse von Nomina helfen. Die Fragen zur Bestimmung der einzelnen Kategorien werden im Anhang (Anlage 2) in einer kompakten Übersicht aufgelistet.

3.2.2 Die Verwendung des Entscheidungsbaums

Wir beginnen ganz oben im Baum und folgen den Pfeilen. Der Ausgangspunkt ist die Frage, ob das Nomen in die Konstruktion „... ist drei Minuten/ Stunden/ Tage/ Monate/ Jahrhunderte/ etc. alt“ eingesetzt werden kann.

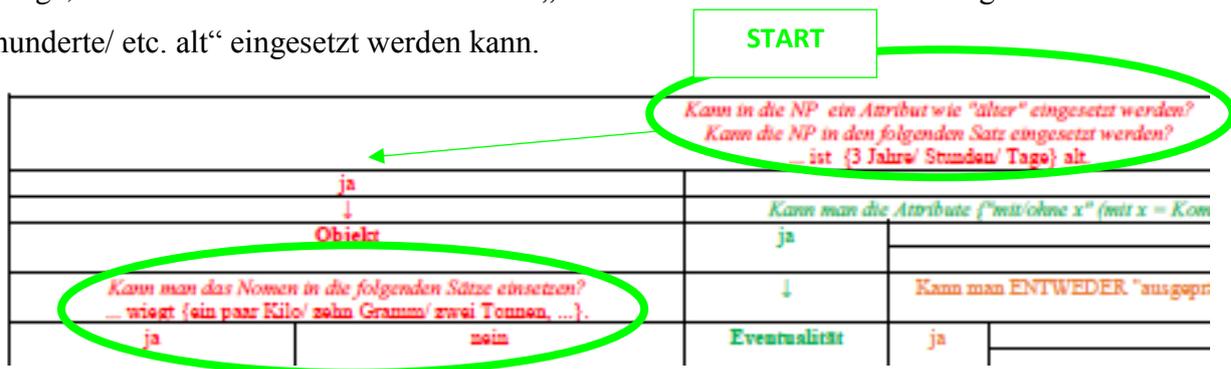


Abbildung 7: Ausgangspunkt des Entscheidungsbaums

Fällt die Antwort positiv aus, so folgt man dem Pfeil unter *ja*. Wir wissen nun, dass es sich um ein Objekt handelt und beantworten die nächste Frage; fällt sie negativ aus, so folgt man dem Pfeil unter *nein*. Diese Prozedur durchläuft man so lange, bis man an einem der Endknoten im Baum ankommt, d.h. dass keine Frage mehr folgt. Dynamische Eventualitäten, statische Eventualitäten und Eigenschaften sind z.B. Endknoten.

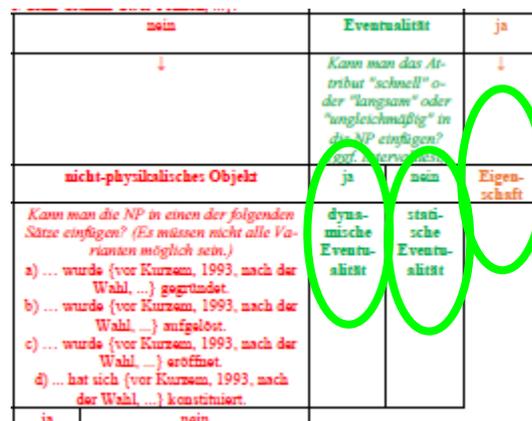


Abbildung 8: Endknoten des Entscheidungsbaums

Die Kategorie, die diesem Endknoten zugeordnet ist, entspricht dem semantischen Typ des Nomens. Im Folgenden werden zwei Beispiele erläutert. Wer sich im Umgang mit dem Baum sicher fühlt, kann bei Abschnitt 3.3 weiterlesen.

Beispiel 1

- (20) Dieser Start und die theatralische Werkqualität werden den "Physikern" wahrscheinlich mehr auswärtige Bühnen öffnen, als sie "Frank V." fand.

Relevant ist in diesem Fall das Subjektnomen *dieser Start*. Für die ersten Fragen im Entscheidungsbaum liefern die Tests negative Ergebnisse:

- (21) a. Dieser (*ältere) Start und die theatralische Werkqualität werden den "Physikern" wahrscheinlich mehr auswärtige Bühnen öffnen, als sie "Frank V." fand.
b. *Dieser Start ist drei Jahre alt.

Aufgrund dessen müssen wir im Baum dem Pfeil unter der Antwort *nein* folgen. Dieser führt uns zu der nächsten Frage. Dabei stellen wir fest, dass es problemlos möglich ist, die Attribute {"mit/ohne x" (mit x = Komitativ, Instrument), "von x begleitet"} in die NP einzufügen:

- (22) a. Dieser Start mit dem bekanntesten Schauspieler des Theaters/ mit einer Glanzleistung/ mit Spezialeffekten wird den "Physikern" wahrscheinlich mehr auswärtige Bühnen öffnen, als sie "Frank V." fand.
b. Dieser von der Presse/ vom Intendanten/ vom Publikum begleitete Start wird den "Physikern" wahrscheinlich mehr auswärtige Bühnen öffnen, als sie "Frank V." fand.

Nun folgen wir dem Pfeil unter der Antwort *ja*. Wir wissen nun, dass es sich um eine Eventualität handelt. Da noch eine Frage folgt, beantworten wir diese.

- (23) Dieser schnelle/ langsame Start wird den "Physikern" wahrscheinlich mehr auswärtige Bühnen öffnen, als sie "Frank V." fand.

Da man „schnell“ einsetzen kann (siehe (23)), folgen wir nun wieder dem Pfeil unter der Antwort *ja*. Dies führt uns zum Endknoten „Dynamische Eventualität“, dem zu annotierenden Typ.

Beispiel 2

- (24) Die Form seines Rumpfes ähnelt der des Weißwals Beluga.
(Die Zeit, 14.09.2012 (online))

Relevant ist in diesem Fall das Subjektnomen *die Form seines Rumpfes*, wobei „Form“ das Kopfnomen und von primärem Interesse ist. Für die ersten Fragen im Entscheidungsbaum liefern die Tests negative Ergebnisse:

- (25) a. Die (?ältere) Form seines Rumpfes ähnelt der des Weißwals Beluga.
b.*Die Form seines Rumpfes ist drei Jahre alt.

Aufgrund dessen müssen wir im Baum dem Pfeil unter der Antwort *nein* folgen. Dieser führt uns zu der nächsten Frage. Dabei stellen wir fest, dass es nicht möglich ist, die Attribute {"mit/ohne x" (mit x = Komitativ, Instrument), "von x begleitet"} in die NP einzufügen:

- (26) a. Die Form seines Rumpfes (*mit Spezialwerkzeugen/ *mit Genmanipulation) ähnelt der des Weißwals Beluga.

- b. Die (*von Schuppen/ *von Delfinen/ *vom Tierschutzbund begleitete) Form seines Rumpfes ähnelt der des Weißwals Beluga.

Deshalb folgen wir dem Pfeil unter *nein*. Dieser führt uns zur nächsten Frage, die ein positives Ergebnis liefert:

- (27) Die immense/ ausgeprägte/ für einen Airbus charakteristische Form seines Rumpfes ähnelt der des Weißwals Beluga.

Wir folgen dem Pfeil unter *ja* und annotieren eine Eigenschaft.

3.3 Der Umgang mit uneindeutigen Fällen

In manchen Fällen ist die Zuordnung einer NP zu einem semantischen Typ nicht ohne Weiteres zu leisten, sondern problematisch. Um den Benutzer dieses Handbuchs für solche Fälle zu sensibilisieren, sollen im Folgenden ein paar Beispiele diskutiert werden, die bei der Annotation Probleme bereitet haben.

- (28) a. In der Ferne erkennt man schemenhaft die offene Flanke, die Bucht, wo sich der Meghna in das Meer ergießt, wo die Zyklone ungebremst die Küste treffen. (Die Zeit, 17.05.2007, Nr. 21)

Das für die Annotation relevante Nomen ist „die Zyklone“. Grundsätzlich scheint ein Zyklon mehrere Typen zuzulassen:

- (29) a. Der Zyklon dauerte fünf Stunden.
b. Der Zyklon bewegte ganze Baumstämme.

Während in (29a) aufgrund der Dauer eine Eventualität vorliegt, liegt in (29b) ein natürliches Objekt vor. Selbst wenn man daher zu einem vermeintlich eindeutigen Ergebnis kommt, muss dieses, gerade weil es so etwas wie dot-Typen gibt (siehe Abschnitt 3.1.2), nochmals kritisch hinterfragt werden. In (28) macht die Annahme eines Ereignisses tatsächlich keinen Sinn: Nur eine in irgendeiner Form aus Materie bestehende Entität kann ein physikalisches Objekt wie *Küste* treffen. Man muss sich daher retrospektiv stets die Frage stellen, ob die annotierte Kategorie im jeweiligen Ko(n)text Sinn macht.

5. Literatur

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.

Asher, N. (2011). *Lexical Meaning in Context. A Web of Words*. Cambridge: Cambridge University Press.

Landis, J.R. und G. G. Koch (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33.1, 159-174.

- Lukassek, J. (2012). Stativ-/eventiv ambige Verben im Deutschen. *Wiener Linguistische Gazette* 76A, 131-152.
- Maienborn, C. (2003). *Die logische Form von Kopula-Sätzen*. Berlin: Akademie-Verlag.
- Maienborn, C. (2005). Eventualities and different things: A reply. *Theoretical Linguistics*, 31, 383–396.
- Maienborn, C. (2005). On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical Linguistics*, 31, 275–316.
- Maienborn, C. (2007). On Davidsonian and Kimian states. In: I. Comorovski and K. von Heusinger (eds.). *Existence: Semantics and Syntax*. Kluwer, Dordrecht. 107–130.
- Maienborn, C. (2015). Events and states. In: Truswell, Robert (ed.): *Handbook of Event Structure*. Oxford: Oxford University Press.
- Maienborn, C. (2017). Konzeptuelle Semantik. In: Sven Staffeldt & Jörg Hagemann (eds.) *Semantiktheorien. Lexikalische Analysen im Vergleich*. Tübingen: Stauffenburg, 151-188.
- Masolo, C., S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, and L. Schneider (2003). *The WonderWeb Library of Foundational Ontologies. Preliminary Report*.
- Moltmann, F. (2013). *Abstract Objects and the Semantics of Natural Language*. Oxford: UP.
- Niles, I. and A. Pease (2001). Toward a Standard Upper Ontology. In: Ch. Welty and B. Smith (eds.): *Formal Ontology in Information Systems. Proceedings of the 2nd International Conference (FOIS-2001)*. New York: ACM Press.
- Pease, A. (2006). Formal representation of concepts: The Suggested Upper Merged Ontology and its use in linguistics. In: A. C. Schalley and D. Zaefferer (eds.): *Ontolinguistics. How Ontological Status Shapes the Linguistic Coding of Concepts*. Berlin, New York: Mouton de Gruyter.
- Pustejovsky, J. (1991). The Generative Lexicon. *Computational Linguistics*, 17(4), 409-441.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: MIT Press.
- Vendler (1967). *Linguistics in Philosophy*. Cornell University Press: Ithaca, New York.