

Diss

Stefan Oppl

7. April 2009

Inhaltsverzeichnis

I	Umsetzung	I
I	Persistierung	3
1.1	Möglichkeiten der Persistenzsicherung	3
1.2	Topic Maps	4
1.2.1	Topics, Subjects, Topic Names und Variants	5
1.2.2	Associations und Roles	8
1.2.3	Occurrences und Datatypes	9
1.2.4	Metamodellierung in Topic Maps	9
1.2.5	Statements und Scopes	13
1.2.6	Reification	14
1.2.7	Merging	15
1.3	Abbildung von Modellen auf Topic Maps	15
1.4	Technische Umsetzung der Persistierung von Modellen	15
1.5	Zusammenfassung	16
	Literaturverzeichnis	17
	Index	17

Abbildungsverzeichnis

I.1	Grundlegende Elemente einer Topic Map	4
I.2	Umfassende Darstellung der Elemente einer Topic Map	6
I.3	Abgrenzung zwischen Subject und Occurrence in Topic Maps	7
I.4	Benennung von Topics	8
I.5	Beziehungen in der Metamodellbildung in Topic Maps	11
I.6	Abbildung von Gültigkeitsbereichen durch Scopes	14

Tabellenverzeichnis

Teil I

Umsetzung

I Persistierung

In den vorangegangenen drei Kapiteln wurde die Umsetzung des eigentlichen Werkzeugs beschrieben. Neben der Unterstützung des Modellierungsvorgangs ist aber auch die persistente Speicherung der erstellten Modelle zum Zwecke der Weiterverarbeitung ein hier zu beleuchtender Aspekt. Auf die Persistierung wirken vor allem zwei der in Kapitel XY identifizierten Anforderungen ein. Zum ersten ist die Nachvollziehbarkeit des Modellierungsvorganges sicherzustellen – dies gilt nicht nur während des Vorgangs selbst, sondern auch danach. Dementsprechend ist sämtliche Information zu persistieren, die zur Wiederherstellung nicht nur des Modells selbst sondern auch der gesamten Modellierungshistorie notwendig ist. Zum zweiten hat die Forderung nach semantischer Offenheit bei der Modellierung auch unmittelbare Auswirkungen auf die Persistierung. Neben dem Modell selbst muss aufgrund dieser Anforderung auch die Bedeutung der verwendeten Modellierungselemente miterfasst und persistiert werden, so dass diese bei der Weiterverarbeitung der Modelle verwendet werden kann.

In diesem Kapitel werden nun aufgrund der eben genannten Forderungen technologische Ansätze identifiziert, beschrieben und schließlich hinsichtlich ihrer Eignung für den konkreten Einsatz beurteilt. Der ausgewählte Ansatz wird im darauf folgenden Abschnitt konzeptuell beschrieben. Die Abbildung der Modelle und der ebenfalls zu persistierenden zusätzlichen Information in ein geeignetes Datenmodell ist Gegenstand des darauf folgenden Abschnitts. Schließlich wird die konkrete technische Umsetzung der Persistierung dargelegt und die dazu notwendigen Software-Module im Detail beschrieben.

I.1 Möglichkeiten der Persistenzsicherung

- Serialisierung von Java-Objekten
- Relationale Datenbanken
- XML Topic Maps

1.2 Topic Maps

Topic Maps (ISO JTC1/SC34/WG3, 2008) sind wie bereits in Abschnitt XY beschrieben ein Mittel zur Abbildung von semantischen Netzen. In Topic Maps können beliebige Daten strukturiert aufbereitet und zueinander in Beziehung gesetzt werden. Die Art der zu repräsentierenden Daten ist dabei irrelevant, eine Topic Map trifft keine Aussage über ein den repräsentierten Daten zugrundeliegendes Begriffssystem (sie ist „ontology-agnostic“ (Vatant, 2004)).

Historisch stammen Topic Maps aus dem Bereich der technischen Repräsentation von Thesauri und Indizes (Pepper, 2000) (Rath, 2003). Aus diesen Bereichen motivieren sich auch die Bausteine einer Topic Map, wenngleich der Verwendung durch diesen Ursprung nicht eingeschränkt wird. Die grundlegenden Elemente einer Topic Map sind „Topics“, „Associations“ und „Occurrences“ (siehe Abbildung 1.1).

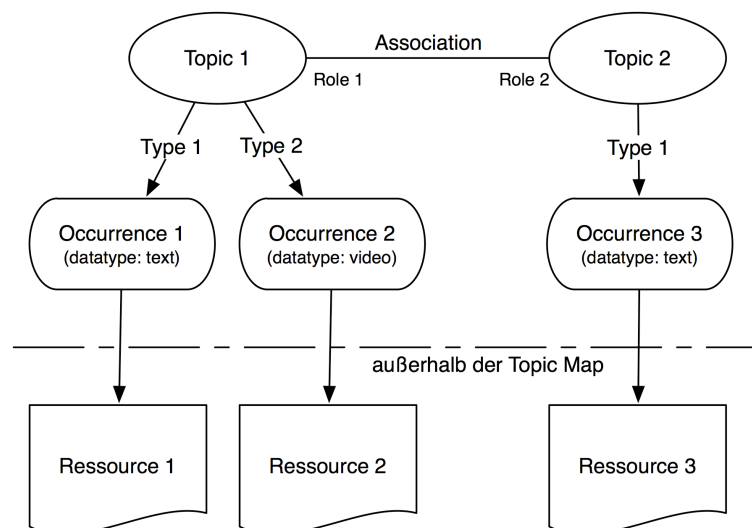


Abbildung 1.1: Grundlegende Elemente einer Topic Map

„Topics“ stellen Begriffe dar und bilden die Knoten des semantischen Netzes. Ein Topic kann beliebige Information darstellen, repräsentiert aber immer genau ein Phänomen der realen Welt (d.h. zu einem Topic muss es eine Entsprechung außerhalb der Topic Maps geben, die beobachtbar oder beschreibbar ist und auf die die modellierende Person Bezug nehmen will ¹). Eine Topic Map ist damit im Sinne

¹ „A subject can be anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever. In particular, it is anything about which the creator of a topic map chooses to discourse.“ (ISO JTC1/SC34/WG3, 2008, S.8)

von Stachowiak (1973) ein diagrammatisches Modell, das einen bestimmten, für den Modellersteller relevanten Ausschnitt der Realität abbildet.

“Associations“ bilden die Beziehungen zwischen Topics ab und stellen damit die Kanten des semantischen Netzes dar. Eine Association verknüpft Topics semantisch miteinander und kann frei mit Bedeutung belegt werden. Die Art der Beziehungen ist also nicht festgelegt und wird wie die Bedeutung der Topics frei gewählt werden. Topics und Associations decken historisch den Bereich der Darstellung von Thesauri ab, in denen Begriffe definiert und zueinander in Beziehung gesetzt werden.

Der zweite historische Ursprung von Topic Maps, die Indizes, werden durch das Konstrukt der „Occurrences“ abgedeckt. Occurrences („Auftreten“) sind Referenzen aus der Topic Map in die reale Welt. Sie setzen die Topics einer Topic Map in Bezug zu beliebiger referenzierbarer Information (z.B. Dokumente). Im Kontext der eben genannten Indizes, kann eine Topic Map als der mit Querverweisen versehene Index eines Buches verstanden werden, in dem durch die Angabe von Seitenzahlen auf den Text des Buches verwiesen wird. Diese Verweise durch Angabe der Seitenzahlen sind in diesem Zusammenhang die Occurrences.

Die Ansammlung von durch Associations verknüpften und mit Occurrences versehenen Topics bilden eine Topic Map. Darüber hinaus kann in Topic Maps jedoch noch weiterführende Information repräsentiert werden (siehe Abbildung 1.2), die Gegenstand der folgenden Abschnitte sein werden.

1.2.1 Topics, Subjects, Topic Names und Variants

Wie oben bereits beschrieben, repräsentiert ein Topic ein Phänomen der realen Welt in einer Topic Map. Dieses Phänomen der realen Welt, das durch das Topic repräsentiert wird, wird als „Subject“ bezeichnet. In einer Topic Map darf es zu einem Subject nur exakt ein Topic geben, umgekehrt kann ein Topic auch nicht mehrere Subjects repräsentieren, die Zuordnung zwischen Subject und Topic ist also eindeutig (bijektiv). Im Topic wird dazu exakt ein „Subject Identifier“ registriert, der auf eine Informationsressource verweist, die das Subject für Menschen eindeutig identifizierbar macht (diese Ressource wird als „Subject Indicator“ bezeichnet). Zusätzlich kann ein „Subject Locator“ angegeben werden, der auf das tatsächlich in der realen Welt vorhandene Subject verweist. In Abgrenzung dazu kann es bei der anderen Brücke zwischen realer Welt und Topic Map, den Occurrences, für jeder Topic beliebig viele Zuordnungen geben. Eine Occurrence referenziert auch auf die reale Welt, zeigt aber dort nicht auf das Subject selbst, sondern auf ein dieses Subject beschreibendes Objekt in der realen Welt. Beispielhaft ist dazu in Abbildung 1.3 dieser Zusammenhang anhand des Topics „Tasse“ dargestellt. Ein anderes Beispiel ist ein Topic „London“, das als Subject die reale Stadt London repräsentiert und dem eine

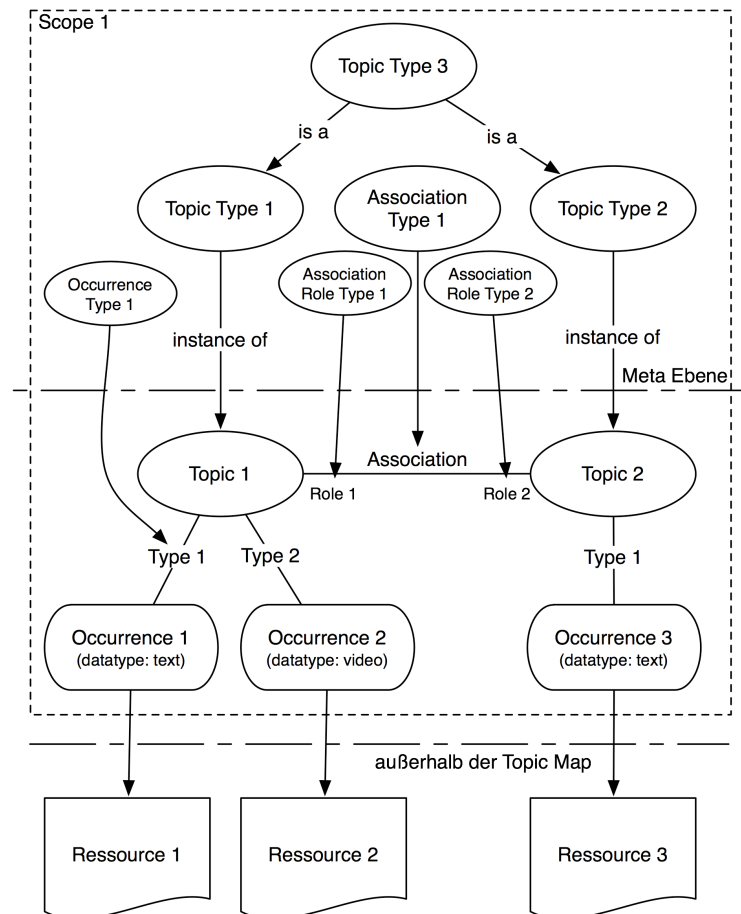


Abbildung 1.2: Umfassende Darstellung der Elemente einer Topic Map

Occurrence zugeordnet werden könnte, die auf eine Landkarte (als in der Realität vorhandene Beschreibung der realen Stadt London) referenziert.

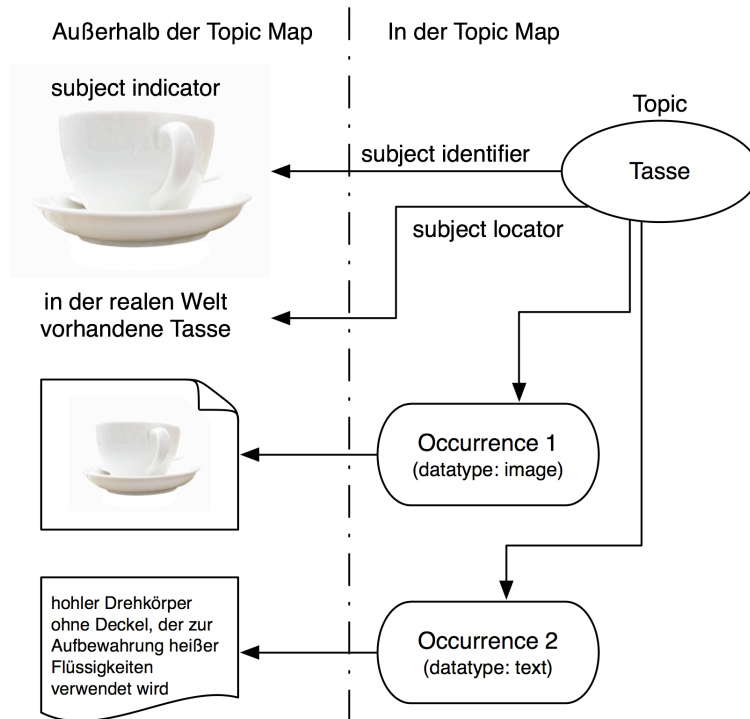


Abbildung 1.3: Abgrenzung zwischen Subject und Occurrence in Topic Maps

Bislang wurde vereinfacht ein Topic immer mit einem direkt zugeordneten Namen dargestellt. In einer Topic Map besitzt ein Topic jedoch keinen eindeutigen Namen. Es wird vielmehr durch seinen Subject Identifier eindeutig gekennzeichnet. Dieser ist jedoch nicht unbedingt für Menschen les- und/oder interpretierbar – der Subject Identifier hat das Ziel, ein Subject für die Verarbeitung durch Software eindeutig zuordnenbar zu machen. Für die Bezeichnung eines Topics in einer für Menschen interpretierbaren Form ist die Verwendung von „Topic Names“ vorgesehen (siehe Abbildung 1.4). Topic Names werden immer textuell angegeben und beschreiben das Subject, das durch das betreffende Topic referenziert wird. Durch einen Topic Name soll das Subject für Menschen erkennbar sein, wobei die Zuordnung nicht notwendigerweise eindeutig sein muss (Beispiel: der Topic Name „Jaguar“ kann ein Fahrzeug oder eine Großkatze bezeichnen und ist dementsprechend ein zulässiger Name für zwei unterschiedliche Topics). Einem Topic können beliebig viele Topic Names zugewiesen werden – es ist so zum Beispiel möglich, eine mehrsprachige Topic Map zu realisieren, in der zu jedem Topic Topic Names in unterschiedlichen Sprachen angegeben werden.

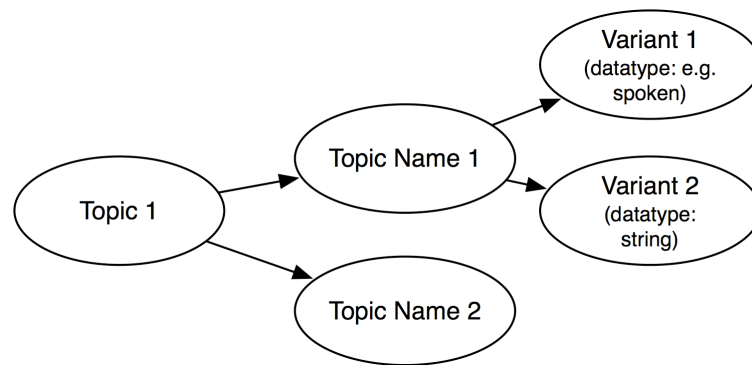


Abbildung 1.4: Benennung von Topics

Als weitere Detaillierungsstufe können zu jedem Topic Name „Variants“ angegeben werden. Wie durch den Namen angedeutet, handelt es sich dabei um Varianten eines Topic Names, die in bestimmten Zusammenhängen oder für gewisse Anwendungszwecke besser geeignet sein können als der eigentlich Topic Name. Ein Beispiel für eine mögliche Variante ist die Angabe einer gesprochenen Version des Topic Names. Ein weiteres Anwendungsgebiet für Varianten ist die im Standard explizit vorgesehene Angabe eines „Sort Name“ (ISO JTC1/SC34/WG3, 2008, S. 18), der es erlauben soll, Topic Maps in eine durch diesen Namen vorgegebene Ordnung bringen zu können. Varianten werden durch die Angabe eines dezidiert dafür gewidmeten Topics in deren Scope (siehe Abschnitt 1.2.5) als Sort Names gekennzeichnet.

1.2.2 Associations und Roles

“Associations“ stellen Verbindungen zwischen den einzelnen Topics einer Topic Map her. Associations haben beliebig viele Endpunkte, mindestens jedoch einen (sind also nicht von vornherein immer binär sondern können auch unär sein oder mehr als zwei Topics verknüpfen). Eine Association enthält wie ein Topic nicht unmittelbar einen für Menschen lesbaren Namen. Diese wird durch ein Topic festgelegt, das die Kategorie festlegt, der die Association zuzuordnen sind (siehe Abschnitt 1.2.4). Diesem Topic kann wiederum mindestens ein Topic Name zugeordnet werden, welcher letztendlich die Benennung der Association festlegt.

Associations werden jedoch nicht direkt mit Topics verknüpft. Um ausdrucksstärkere Verknüpfungen realisieren zu können, agieren „Roles“ als Verknüpfung zwischen Association und den betreffenden Topics. Roles legen die „Rolle“ – also die Bedeutung – eines Topics in exakt der betrachteten Association fest. Diese Bedeutung kann generisch sein und zum Beispiel dazu verwendet werden, die per se ungerichteten Associations unabhängig von ihrer konkreten Bedeutung mit einer Richtung

zu versehen (zum Beispiel durch die Zuordnung von Roles „Anfang“ und „Ende“) aber auch um die Beziehung semantisch anzureichern (zum Beispiel durch die Zuordnung von Roles „Verantwortlicher“, „Ausführender“ und „Prozessschritt“ in einer Association „durchzuführen“). Die Anzahl der in einer Association referenzierten Roles gibt damit auch die Kardinalität der Association (also die Anzahl ihrer Endpunkte) an. Aus den konkreten Roles wird dann auf die Topics verwiesen, die diese Roles einnehmen bzw. „spielen“ (tatsächlich heißt die betreffende Eigenschaft einer Role „player“). Genau wie Associations werden Roles nicht direkt benannt sondern über ein Topic, das ihre Kategorie bestimmt, mit einer Benennung versehen (Detail wiederum in Abschnitt 1.2.4).

1.2.3 Occurrences und Datatypes

Wie zu Beginn dieses Abschnitts bereits beschrieben und in den Erläuterungen zur Thematik der Subjects (siehe Abschnitt 1.2.1) angedeutet, bilden „Occurrences“ die Brücke aus der Topic Map in die reale Welt, indem sie auf Ressourcen referenzieren, die in einem beliebigen Zusammenhang mit den jeweiligen Topics stehen. Ein Topic kann beliebig viele Occurrences haben. Anders als bei Associations existieren für Occurrences keine Roles (was auch nur bedingt sinnvoll wäre, da jede Occurrence nur zu exakt einem Topic gehören kann). Die Bedeutung der Occurrence für das Topic kann wie bei Associations über die Kategorie der Occurrence festgelegt werden, die wiederum durch ein separates Topic repräsentiert wird (siehe dazu auch Abschnitt 1.2.4). Beispielsweise kann eine Occurrence zur Kategorie „Karte“ gehören und so angeben, dass die so klassifizierte Occurrence zum Topic „London“ auf eine Karte des Stadtgebiets verweist.

Zusätzlich zu der Kategorie wird in einer Occurrence auch der Datentyp der Information angegeben, in dem die referenzierte Information vorliegt. Dabei können beliebige URI (Uniform Resource Identifiers²) verwendet werden. Da URIs beliebigen Inhalt haben können, wäre es in obigen Beispiel möglich durch den Datentyp einer Occurrence festzulegen, ob es sich bei der Karte um eine Rastergrafik oder eine Vektorgrafik handelt und so Information über deren möglich Einsatzgebiete zu einzubetten.

1.2.4 Metamodellierung in Topic Maps

Wie oben bereits mehrmals angedeutet kann in einer Topic Map neben den eigentlichen zu repräsentierenden Informationen (Topics, Associations, Roles und Occurrences) auch Information über die Topic Map selbst eingebettet werden (neben

²wie in RFC 3986 definiert und unter <http://www.ietf.org/rfc/rfc3986.txt> abzurufen

dem Model also auch das Meta-Modell abgebildet werden kann). Die Information umfasst Angaben über die in der jeweiligen Topic Map existierenden Kategorien von Topics, Topic Names, Associations, Roles und Occurrences. Hinsichtlich der Repräsentation dieser Information sind zwei Ansätze zu unterscheiden, von denen der erste bei Kategorieangaben von Topics zum Einsatz kommt, der andere bei Kategorieangaben jeder anderen Art von Information. Allen Kategorien ist gemein, dass sie selbst wiederum als Topics repräsentiert werden und auch als solche verwendet werden können. Es ist also möglich, zu einem Topic, das als Kategorie verwendet wird, selbst wiederum eine Kategorie anzugeben, wodurch die Einführung beliebig vieler Meta-Ebenen möglich ist. Außerdem können als Kategorien verwendete Topics ebenfalls wieder mit Associations verknüpft und mit Occurrences versehen werden. Hinsichtlich der Nomenklatur ist noch darauf hinzuweisen, dass Kategorien im Allgemeinen als „Types“ bezeichnet werden, man also von „Topic Types“, „Topic Name Types“, „Association Types“, „Role Types“ und „Occurrence Types“ spricht.

Topic Types

Topic Types werden in einer Topic Map durch eine spezielle, im Standard festgelegte Association definiert. Soll einem Topic ein Type zugewiesen werden, muss eine Association der Kategorie „type-instance“ eingefügt werden, bei der das Topic selbst die Rolle „instance“ einnimmt und dem Topic, das die Kategorie repräsentiert, die Rolle „type“ zugewiesen wird. Diese Beziehung entspricht einer Konkretisierung einer (abstrakten) Kategorie oder Klasse von Topics auf eine bestimmte Instanz, die die Merkmale dieser Klasse trägt. In Abbildung 1.5 besteht eine type-instance-Beziehung (dort als „instance-of“ bezeichnet) zwischen der Kategorie „VW Golf“ und der konkreten Instanz „SR-174 AU“ (also einem Topic, bei dem das amtliche Kennzeichen als Topic Name verwendet wurde). „VW Golf“ fungiert hier also als Topic Type, wobei es selbst ein Topic ist, das sich durch nichts als die eingenommene „type“-Rolle von einem anderen Topic unterscheidet und dementsprechend behandelt werden kann.

Einem Topic können beliebig viele Topic Types zugewiesen werden, indem es in mehr als einer Association die Rolle „instance“ einnimmt. Es wird so als mehreren Kategorien zugehörig gekennzeichnet. Umgekehrt kann ein Topic Type mehr als einem Topic zugewiesen werden, indem das betreffende den Topic Type repräsentierende Topic die Rolle „type“ mehrfach einnimmt.

Auch ist es möglich, Hierarchien von Types zu bilden, in dem einem Topic, das als Topic Type fungiert, selbst wieder ein Topic Type zugewiesen wird. Dieses Konstrukt ist jedoch mit Vorsicht zu gebrauchen, da zur Abbildung der Struktur einer Domäne ein semantisch ähnliches, bei näherer Betrachtung aber eine unterschiedliche Bedeutung tragendes Konstrukt zum Einsatz kommt. Grundsätzlich muss un-

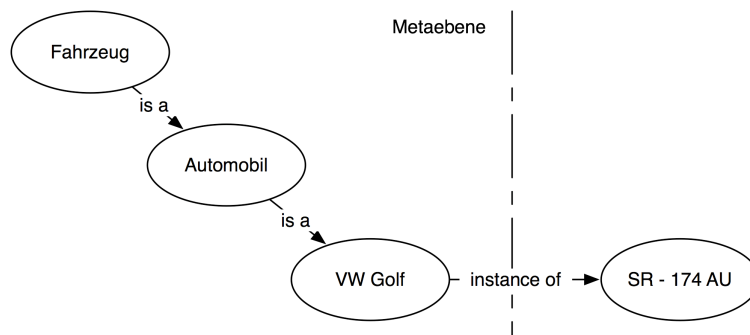


Abbildung 1.5: Beziehungen in der Metamodellbildung in Topic Maps

terschieden werden, ob ein Topic eine konkrete Instanz eines anderen ist oder lediglich eine Spezialisierung darstellt. Im ersteren Fall kommt eine „type-instance“ zum Einsatz, das übergeordnete Topic befindet sich semantisch auf einer anderen, abstrakteren Ebene und stellt eine Kategorie (also einen Topic Type) dar. Im Falle einer Spezialisierung kommt eine „supertype-subtype“-Association zum Einsatz, deren Rollen „supertype“ vom übergeordneten, allgemeineren bzw. „subtype“ vom untergeordneten, spezielleren Topic eingenommen wird. Hier befinden sich beide Topics semantisch auf einer Ebene, keines ist abstrakter als das andere. Der Unterschied liegt vielmehr in der mehr oder weniger konkreten Festlegung der durch die Topics repräsentierten Subjects. So ist wie in Abbildung 1.5 dargestellt das Topic „VW Golf“ ein Subtype des Topics „Automobil“ welches wiederum ein Subtype des Topics „Fahrzeug“ ist. Hier ist erkennbar, dass „Automobil“ insofern eine Spezialisierung von „Fahrzeug“ ist, als dass es im Allgemeinen motorisiert ist und vier Räder besitzt. „VW Golf“ ist wiederum eine Spezialisierung von Fahrzeug, die hinsichtlich der Form der Karosserie, der Anzahl der Türen und anderen Merkmalen mehr spezialisiert. Zwischen keinem der Topics findet jedoch eine Konkretisierung in dem Sinne statt, als dass auf der untergeordneten Seite von einem konkreten, real existierenden Fahrzeug die Rede wäre – dazu ist die „type-instance“-Beziehung zu verwenden. Die „subtype-supertype“-Beziehung ist transitiv, d.h. dass ein „VW Golf“ nicht nur ein „Automobil“ ist, sondern auch ein „Fahrzeug“. Für „type-instance“-Beziehungen ist diese Eigenschaft nicht gegeben.

Andere Types

Bei allen anderen Types (konkret also Topic Name Types, Association Types, Role Types und Occurrence Types) wird die Kategorie nicht durch eine separate Association abgebildet sondern durch eine im jeweiligen Informationselement enthaltene Referenz auf ein Topic, das als Type fungiert. Die Darstellung der Kategorie-

Information ist damit nicht so explizit wie bei Topic Types, wo sich direkt in der Repräsentation der eigentlichen Nutzinformation niederschlägt. Auch semantisch ist sind die hier behandelten Types gegenüber Topic Types insofern eingeschränkt, als das jedem Element exakt ein Type zugeordnet sein muss (der Type kann also nicht leer sein, auch können nicht mehrere Types zugeordnet werden). Wie oben bereits erwähnt sind Topic Types hier flexibler, einem Topic können beliebig viele oder auch keine Topic Types zugeordnet werden. Dies ist aber nur vordergründig eine Einschränkung. Topics dürfen wie oben beschrieben für jedes Subjekt nur einmal existieren. Hat aber ein Subject und damit ein Topic in unterschiedlichen Domänen unterschiedliche Bedeutungen, muss dies über mehrere Topic Types (in Verbindung mit Scopes, siehe Abschnitt 1.2.5) abgebildet werden. Alle anderen Informationskategorien in der Topic Map unterliegen nicht dieser Eineindeutigkeitsregel und können bzw. müssen, sollten sie unterschiedlichen Kategorien zuzuordnen sein, auch mehrfach vorhanden sein. Eine Assoziation, die einen anderen Namen trägt (also einer anderen Kategorie angehört) ist beispielsweise nicht identisch mit der ursprünglichen Assoziation, deren Name ebenfalls bereits durch die Zuordnung zu einer Kategorie festgelegt wurde.

Modellieren von Einschränkungen

Der Topic Map Standard erlaubt zwar die Angabe von Metamodellelementen (Types), ermöglicht es aber nicht Regeln anzugeben, anhand derer der semantisch korrekte Aufbau einer Topic Map geprüft werden. Es ist beispielsweise möglich, einen Association-Type „hat Mitglieder“ zu definieren, der mittels den Roles „Organisationseinheit“ und „Mitarbeiter“ die Zuordnung von Mitarbeitern zu den Organisationseinheiten eines Unternehmens zuzuordnen. Es ist jedoch in der Topic Map nicht möglich zu spezifizieren, dass beispielsweise mindestens drei Mitarbeiter zugeordnet werden müssen oder dass es in dieser Beziehung nur eine Organisationseinheit geben darf. Weiters kann nicht spezifiziert werden, durch Topics welchen Types die jeweiligen Roles eingenommen werden dürfen – beispielsweise ist eine Zuordnung von Produktionsmitteln in der Role „Mitarbeiter“ zulässig bzw. kann sie nicht als unzulässig gekennzeichnet werden.

Ist eine derartige semantische Einschränkung bei der Topic Map Erstellung und die Einführung verbindlicher Strukturvorgaben notwendig, so muss dies außerhalb der Topic Map oder durch externe Interpretation spezifischer Topic Map Elemente geschehen. In ersterem Fall kann die noch finale Zustand vorliegende TMCL (Topic Map Constraint Language), eine XML-basierte Regelsprache zur Einschränkung der Repräsentationsmöglichkeiten in einer Topic Map, verwendet werden. Im zweiten Fall können die Metamodell-Elemente (also alle Topics, die als Types verwendet werden) durch zusätzliche Associations verknüpft werden, die semantisch

so interpretiert werden, dass sie eine zulässige Kombination von Topics der jeweiligen Kategorie anzeigen.

1.2.5 Statements und Scopes

Topic Maps bieten die Möglichkeit, Gültigkeitsbereiche für die in ihnen abgebildeten Informationen zu spezifizieren. Ein Gültigkeitsbereich definiert, in welchem Kontext eine Information gültig ist. Außerhalb dieses Kontext kann über die Gültigkeit keine Aussage getroffen werden. Der Gültigkeitsbereich wird als „Scope“ bezeichnet. Ein Scope kann für jedes „Statement“ in der Topic Map gesetzt werden. Statements sind alle „Aussagen“ über Topics, die in der Topic Map abgebildet werden, nicht aber Topics selbst. Als Statement werden Topic Names, Associations und Occurrences betrachtet. Roles und Variants besitzen keinen Scope, da sie keine direkte Aussage über Topics treffen, sondern nur im Zusammenhang mit Associations bzw. Topic Names existieren, deren Scope sich quasi auf sie vererbt.

Ein Scope für ein Statement wird durch die Angabe eines oder mehrerer Topics festgelegt. Wird kein Topic angegeben, so gilt der „unconstrained scope“, das Statement ist unbeschränkt gültig. Topics zur Abbildung von Scopes können explizit angelegt werden (z.B. Topics „Deutsch“ und „Englisch“, die Sprach-Scopes ermöglichen, Topics „Tierwelt“ und „Transportwesen“ zur Domänenabgrenzung – siehe Abbildung 1.6), es ist jedoch grundsätzlich auch möglich, die Gesamtheit der Topics einer Domäne zur Definition eines betreffenden Scopes heranzuziehen (also alle in der Topic Map vorhandenen Topics, die Tiere repräsentieren, als Scope zu verwenden, um den Gültigkeitsbereich „Tierwelt“ abzubilden). Obwohl der Topic Map Standard hier explizit offen bleibt, erscheint jedoch erstere Variante hinsichtlich der Verwaltbarkeit aber auch der semantischen Vollständigkeit wegen als ratsamer (das Konzept „Tierwelt“ käme ansonsten z.B. nicht notwendigerweise vor, sondern wäre nur implizit vorhanden, was eine Auswertung schwierig macht).

Wird mehr als ein Topic angegeben, so bilden alle Topics gemeinsam den Kontext, indem das Statement gültig ist. Bei der Auswertung des Gültigkeitsbereichs müssen also alle angegebenen Topics zutreffen, damit ein Statement gültig ist. Ein Statement dessen Scope die Topics „Tierwelt“ und „Deutsch“ enthält, ist also nur gültig, wenn beide Topics zutreffen (also z.B. zur Filterung ausgewählt wurden). Die gültigen Statements sind also die Schnittmenge jener der Statements, die im Scope „Tierwelt“ und im Scope „Deutsch“ gültig sind (siehe Abbildung 1.6). Die Angabe eines Scopes zu einem Statement, das in beiden Scopes (unabhängig voneinander gesehen) gültig sein soll, ist nicht direkt möglich. In diesem Fall muss das Statement zweimal jeweils unter Angabe eines der beiden Scopes eingefügt werden.

Für Topics selbst kann kein Scope angegeben werden. Das ist hinsichtlich des To-

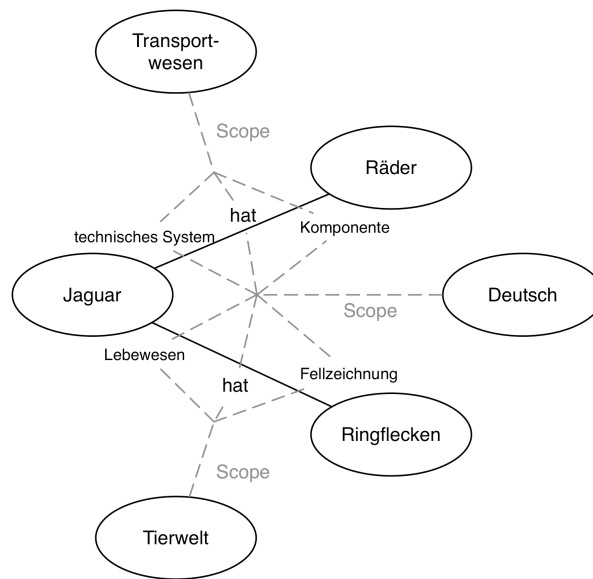


Abbildung 1.6: Abbildung von Gültigkeitsbereichen durch Scopes

pic Maps zugrunde liegenden Konzepts auch nicht sinnvoll, da Topics immer ein Phänomen der realen Welt, das Subject, repräsentieren, das selbst immer vorhanden ist und dessen Gültigkeit bzw. Existenz nicht von anderen Rahmenbedingungen wie anderen Subjects abhängt. Dementsprechend existiert ein Topic immer, lediglich seine Bezeichnung, die Beziehungen zu anderen Topics oder seine Occurrences können abhängig vom einem Scope unterschiedlich sein.

1.2.6 Reification

Wie bereits zu Beginn angeführt, können in einer Topic Map beliebige Inhalte abgebildet werden, jegliche Phänomene der realen Welt können durch Topics repräsentiert werden. Konsequenterweise können nun auch Elemente der Topic Map selbst oder sogar die Topic Map ansich durch ein Topic dargestellt werden. Eine derartige selbstbezügliche Abbildung wird als „Reification“ bezeichnet. Der Ausgangspunkt für eine Reification kann ein Statement sein oder eine Topic Map selbst. Topics können nicht verwendet werden, da ein sie repräsentierendes Topic semantisch äquivalent mit dem Ausgangspunkt wäre und damit zwei Topics vorhanden wären, die das gleiche Subject referenzieren. Nachdem die Abbildung zwischen Subject und Topic eineindeutig sein muss, ist ein derartiges Konstrukt nicht erlaubt.

In Abgrenzung zur Types (Association Types, Occurrence Types, ...) repräsentiert ein reifizierendes Topic nicht die Kategorie eines Elements sondern das konkrete

Element selbst. Es ist so möglich, einem konkreten Element wie einer Association zusätzliche Information hinzuzufügen (etwa eine Occurrence) oder auch eine bestehende Topic Map als Ganzes zu kapseln und als solche durch das sie reifizierende Topic Bezug zu nehmen.

1.2.7 Merging

Unter „Merging“ versteht man die Vereinigung zweier voneinander getrennter Topic Maps zu einer gemeinsamen Map. Dabei ist vor allem die Eineindeutigkeitsregel zu beachten, für ein Subject darf also in der resultierenden Topic Map nur ein Topic existieren. Strukturell nicht kritisch, jedoch die Verwendbarkeit einschränkend sind mehrfach auftretende semantisch identische Statements. Soweit möglich, sollten auch diese Duplikate im Merging-Prozess entfernt werden.

Der Topic Map Standard definiert für jede Art von Element Regeln, anhand der festgestellt werden kann, ob zwei Elemente dieser Art identisch sind oder nicht. Ausgangspunkt sind immer die Topics, wobei beim Vergleich ausschließlich vom abgebildeten Subject ausgegangen werden muss. Sind die Subjects identisch, sind im Wesentlichen auch die beiden Topics identisch und können durch eine gemeinsames Topic ersetzt werden. Auf Basis der vereinigten Topic Menge werden nun die enthaltenen Statements verglichen. Damit Statements als identisch erkannt werden, müssen nicht nur ihr eigentlicher Inhalt sondern auch ihre Kategorie (Type), ihr Gültigkeitsbereich (Scope) und das/die ihnen zugeordnete(n) anderen Element(e) identisch sein. Eine Role ist also nur dann mit einer anderen Role identisch, wenn ihr Type, ihr Scope, die Association, der sie angehört und das referenzierte Topic identisch ist. Daraus folgt, das in der Merging-Reihenfolge zuerst die Topics, dann die eigenständigen Statements wie Topic Names, Associations und Occurrences und letztendlich die abgängigen Statements wie Roles und Variants behandelt werden müssen.

1.3 Abbildung von Modellen auf Topic Maps

-> DA Matthias

1.4 Technische Umsetzung der Persistierung von Modellen

Topic Map Engine Persistence Layer

I.5 Zusammenfassung

Literaturverzeichnis

ISO JTC1/SC34/WG3 (2008). Information Technology - Topic Maps - Part 2: Data Model. International Standard 13250-2, ISO/IEC.

Pepper, S. (2000). The tao of topic maps. In *Proceedings of XML Europe*.

Rath, H. (2003). *The Topic Maps Handbook*, empolis GmbH.

Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Springer Wien.

Vatant, B. (2004). Ontology-driven Topic Maps. In *Proceedings of XML Europe 2004*, Amsterdam.