

# *PrimaVera* Working Paper Series



UNIVERSITEIT VAN AMSTERDAM

*PrimaVera* Working Paper 2006-03

## **Topic Maps uprooted**

Pieter Wisse

January 2006

Category: academic

University of Amsterdam  
Department of Information Management  
Roetersstraat 11  
1018 WB Amsterdam  
<http://primavera.fee.uva.nl>

Copyright ©2006 by the Universiteit van Amsterdam

All rights reserved. No part of this article may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the authors.

## **Topic Maps uprooted**

Pieter Wisse

**Abstract:** Topic Maps is an International Standard, aiming at a “notation for interchangeably representing information about the structure of information resources.”[4] An informal introduction by S. Pepper goes beyond an orientation at resources, traditionally called documentation. He presents Topic Maps as a “standard for describing knowledge structures and associating them with information resources.” He holds it up as “the Global Positioning System of the information universe,” as Topic Maps permits “to encode arbitrarily complex knowledge structures.”[9]

Topic Maps, is it really the silver bullet for the information society? This paper conducts a critical appraisal. My review is not comprehensive, though. I’ve especially attempted a thorough analysis of conceptual foundations, with a special concern for standardized rules and recommendations for map construction. As theory, or metamodel, in the final section Topic Maps is also briefly compared to Metapattern which is the author’s design for controlling requisite variety in information modeling. Some references on both Topic Maps and Metapattern are included at the end of this paper.

**Keywords:** Topic Maps, metapattern, scope, context, information modeling, conceptual modeling.

**About the author:** Dr ir Pieter Wisse ([www.wisse.cc](http://www.wisse.cc)) is informatiekundig ontwerper en oprichter van Information Dynamics, een onafhankelijk advies-/ontwerpbureau voor onderzoek & ontwikkeling van (digitale) infrastructuur voor informatievoorziening met hoge variëteit. Hij is tevens verbonden aan PrimaVera, het onderzoekprogramma in informatiemanagement aan de Universiteit van Amsterdam.

**INDEX**

---

<b>1. Obstacle: different professional perspectives .....</b>	<b>4</b>
<b>2. Retrofitting Topic Maps .....</b>	<b>4</b>
<b>3. Welcome to the community of conceptual modeling.....</b>	<b>7</b>
<b>4. Intermediary remarks.....</b>	<b>10</b>
<b>5. Heritage of names.....</b>	<b>11</b>
<b>6. After Babel: scope .....</b>	<b>13</b>
<b>7. Different perspectives on behavior .....</b>	<b>16</b>
<b>8. The case for requisite variety .....</b>	<b>18</b>
<b>Literature .....</b>	<b>20</b>

## **1. Obstacle: different professional perspectives**

---

For the sake of exposition, I assume two professional communities with separate roots. Then, on the one hand the principles of Topic Maps are readily understandable for members of the traditional information professions, such as archivists and librarians (A&L). For that is precisely the community it originated from ... to cater to initially. As I will show, even as opportunities for wider application are now claimed, its original perspective of documentation management still pervades Topic Maps.

On the other hand, it is already a problem getting first of all Topic Maps more generally understood. Its proponents, that is, mostly professionals with roots as archivists and/or librarians, seem largely unaware of what quite independently has emerged as a different professional tradition. For since about the 1950s, a community developed that is now much greater in number. Indeed, most people active in all sorts of sub-disciplines of modern information management do not, at least when they're honest, have a clue what the millennia-old tradition of documentation management entails. However regretful, that's just how it is. By formal education and practical work experience, their dominant perspective concerns digital information technology (IT), applied for management & control, workflow, games, etcetera. Wasteful as it is, many proven principles and practices of documentation management are being reinvented.

So, starting out from a different perspective IT professionals easily miss the points Topic Maps is making. My concern here in analyzing Topic Maps is that myopia also works the other way around. Topic Maps, too, displays several attempts at reinvention. It fails to apply a wider orientation to information management, Not recognizing achievements elsewhere, i.e. in the community of 'digital technology workers,' combined with making claims in wordings unfamiliar to members of the new(er) information professions, all tends to consolidate Topic Maps' isolation (rather than overcoming it).

As digital technology is widely used now for libraries and archives, theoretical, methodological etc. integration it is only natural. At the current stage of — lack of — convergence, though, it helps to be aware of separate intellectual origins.

## **2. Retrofitting Topic Maps**

---

I may seem to move away from convergence by starting to explain Topic Maps in terms of its roots, i.e. the community of practice for documentation management. However, I believe to narrow the conceptual gap between documentation workers versus digital technology workers to bridgeable proportions, when the theory of Topic Maps is first of all more clearly recognized as the extension from within the tradition of A&L that it really is. Indeed, Topic Maps is usually presented as augmenting a thesaurus.[9] And a thesaurus already augments a straightforward index.

For a clear idea of roots, I take my first bearings at index. Let an index merely be a list of terms, either simple or compound terms, each followed by reference(s) to the location(s) where the corresponding subject is treated at some length.

Now please note that, in the A&L community of documentation management, *subject* is a highly privileged term. For the moment, I'll leave it. As a fundamental approach, I first suggest a semiotic analysis of index.

C.S. Peirce (1839-1914) argued that semiotics involves an irreducible triadic relationship; a sign is one of three constituting elements, object and interpretant being the other two elements. Object? An object is what the sign mediates a particular person to believe, through an interpretant, about reality.[7, 16]

I admit it is difficult not to get confused. What the A&L community practices for its methods etc., is abstraction from individual beliefs, or interpretants. And, as I explained, it calls an object a subject. It *still* calls it that, actually, for what is now more generally referred to as object was once more generally known as ... subject. So, that's how old the tradition of documentation management really is!

Subjects reside in, let me state briefly, reality at large. And for the time being neglecting Peirce's postulate that sign essentially mediates, a sign may be simply taken as whatever immediate description of reality. A sign doesn't describe all of reality, though. What 'part' it does is, completing the necessarily circular argument around basic assumptions, a subject. Please note I'm still intending subject to mean what documentation professionals mean by it.

How does this help understanding what an index is? No, I'm not trying to distract attention by mentioning that an academic tradition in Peirce scholarship has been established, i.e. yet another community of practice. Through mistaken fundamentalism, its members will undoubtedly object (sic!) to index figuring as my ground for retrofitting Topic Maps. Index, according to Peirce, is a kind of sign, and my treatment of course doesn't do full justice to his definition. Well, here is a different context. What about Peirce's concept of index not doing justice, just here? I'll certainly have something to say about context, and situation, later on.

All right, then, an index. In fact, it presupposes a document. Now, our inability to comprehend reality as an unarticulated whole — it is first of all highly impractical to do so, I would say — often leads us, instead, to assume reality as a set of subjects. What does this fragmentation into — possible — subjects do to how we may consider a particular document? Precisely, it can be divided into, say, document parts (and so on, up to a practical limit). From a semiotic perspective each document part, then, is a sign that is about one or more subjects. I want to stress that between reality and document, the former 'holds' subjects and the latter signs. Figure 1 outlines the argument so far.

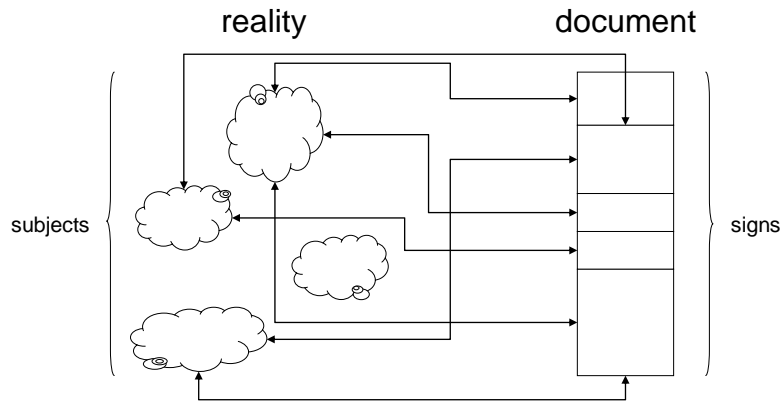


Figure 1: The subjects of reality relate to the signs of document, vice versa.

Adding an index to a document, I suggest, is simply repeating the procedure. A shift occurs, though. The index's reality is no longer reality at large, but more specifically the document, only. From the point of view of the index, a document part is now — treated as — a particular subject. An index entry is considered a sign in this connection (also read: in this situation; more about situation later). Figure 2 sketches such cascading referrals. What comes out nicely in figure 2, is that document locations/parts where subjects are treated, are often only vaguely indicated through index entries, too.

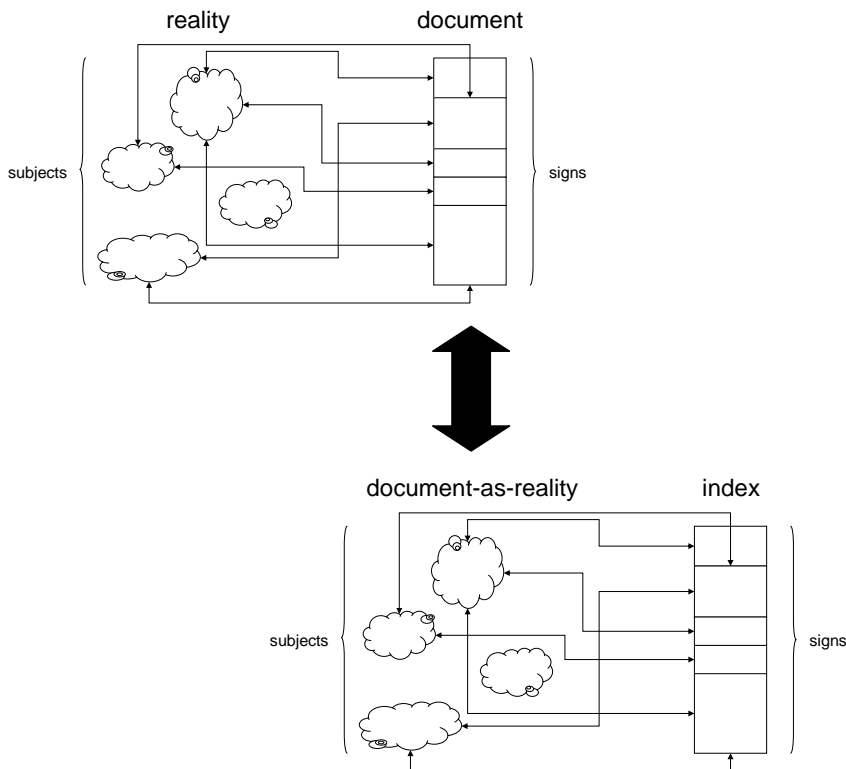


Figure 2: Procedural shift, or subjects/parts of documents relate to signs of index, vice versa.

It is important to appreciate the significance of the threefold distinction between reality, document, and index. What is known as subject classification should be understood as involving

1. a reality at large made up — thought up, really — of subjects,
2. documents where parts describe subjects, and
3. indices where entries label subjects that refer to documents parts describing them.

The formal distinction between description and label is of course weak, non-existent, in fact, as will also become clearer later on. For now, I have arrived at the point where I can already put the Maps of Topic Maps in perspective.

What qualifies as a map? When I first orient myself at a straightforward index, this question has an easy answer. An index provides a map of the document.

I emphasize that an index is a document map because document in its turn, and above all to start with, is considered a map of reality. However implicit, that's really what makes an index work. When the document-as-map-of-reality seems too cumbersome to consult directly, the index-as-map-of-document offers more manageable directions. It essentially makes an index a map of a map.

This indirect aspect of index is habitually overlooked, but nonetheless constitutes its very ground. Anyway, an index helps to navigate the terrain a document constitutes. Therefore an index certainly deserves to be awarded map-status, too.

It doesn't matter that an index is a list, only. As I said, it maps a document, period. So, in this respect Topic Maps is not different from an index. Also being a map, apparently it is a different kind of map. What, then, makes Topic Maps different? In the next section I'll explain how Topic Maps no longer fundamentally revolves around one or more documents, as separate indices do. But for Topic Maps, departing from a traditional index amounts to becoming similar to what digital technology professionals have been doing all along. As I already argued, such convergence is a normal development.

### **3. Welcome to the community of conceptual modeling**

---

A list-based index has limited document mapping potential, leave alone for ever-growing document collections. With broadening scope in space and/or time, variety (also read: ambiguity) manifests itself. Simply a label, as an index for a separate document holds, can no longer be counted upon to efficiently direct the 'documentation traveler' to relevant destinations.

Catalogues come in several shapes and sizes. The general idea is that relationships are introduced for extra structural expression. Such relationships may be prefabricated, as with Dewey's Decimal Classification where decimal notation allows broadening, respectively narrowing of subject areas. Or both labels and relationships may be determined for a particular catalogue, i.e. as a map for a

particular collection: a thesaurus. So, compared to an index, a thesaurus incorporates more structure. It includes suggestions for — labels for — related subject descriptions.

But additional structure is only the effect of an added orientation, one that soon takes precedence. For it should be noted that relations among subject labels, that is, between the entries of a thesaurus, are established not so much by analyzing a particular document. On the contrary, the primary purpose of relating labels is that relevant parts from different documents are all selected for consideration, even when actual labels might differ.

The map character of a thesaurus acquires a dual nature through subject labels being related within its own domain. It remains a map of the managed collection of documents. But now an effort is also undertaken to *directly* map reality. That is, the relationships a thesaurus establishes between subject labels necessarily abstract from separate documents.

How a thesaurus traditionally maps reality still remains crude. It happens on account of the limited repertoire of relationships. The essential step, though, is that especially the structural quality of the thesaurus-as-map is now also directly oriented at reality (rather than predominantly indirectly, as still is the case for an index). Mathematician-turned-librarian S.R. Ranganatan designed a major contribution with his Colon Classification.[10] Widely unrecognized, regrettably, remains especially his construct of recursive application of facets; recursion is flexible hierarchy.

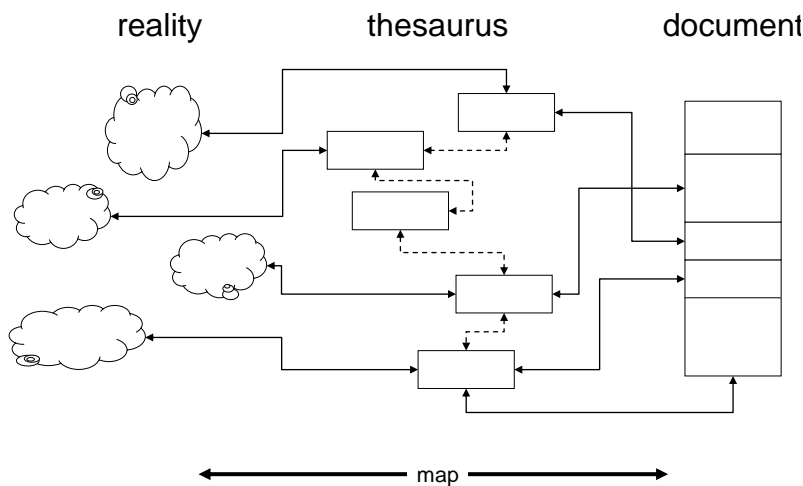


Figure 3: A thesaurus effectively combines two maps directly, i.e. of reality at large and of documentation collection.

Figure 3 sketches how priority for controlling documentation collections moves from an index as a document map to the thesaurus as a map of reality. Mapping individual documents is now performed subsequently, always within the thesaurus' structural possibilities.

Is my picture of this order too one-sided? Yes, of course, somewhat. Any new acquisition to — or deletion from, for that matter — a collection should potentially influence the thesaurus as dually



oriented map. But my primary aim here is to place Topic Maps in perspective. It helps when already a thesaurus is understood as a direct map of reality. And already for a thesaurus, not the documents but reality at large gains priority for orientation to information modeling.

What I've therefore deliberately left out, in figure 3, is how a document of course still maps reality directly, as shown in figure 1. For what figure 3 underlines is how a thesaurus now also directly maps reality (and indirectly, as an index does, by referring from an entry to one or more documents, etc.).

So, an index already is a map. In addition, a thesaurus already expresses structure through explicit relationships between subject labels. As topic, at least according to the International Standard for Topic Maps, is another label for ... subject label, what differentiates it?

Starting from a thesaurus, the answer is: more power for structural expression. I would therefore argue that Topic Maps is a development by degree at the most, and not at all a radical departure. Topic Maps' enhanced potential comes, basically, from

1. custom-typing the subject labels and
2. custom-typing the relationships between subject labels.

Essentially reducing building blocks to variable subject label (also read: topic) and variable relationship (also read: association), whatever may be represented by graphs in mathematics can be at least theoretically accommodated with Topic Maps.

Referring to graph theory should make it even more clear that Topic Maps has a lot in common with all kinds of earlier developments elsewhere. For another example, I fail to see the difference with the basic building blocks of Entity-Relationship Modeling, i.e. under the constraint that an attribute is always a related entity, which can be rephrased as a theory of Entity Maps.

Indeed, the label Topic Maps had me confused. At face value, I thought it was a method for mapping topics. Little did I know what was meant by topic. I've now discovered my initial interpretation was mistaken (although I don't feel helped by the International Standard's sloppiness; it mixes fundamental semiotic categories/dimensions; but I now have to admit that the label 'information modeling' might be equally confusing; at least I avoid 'knowledge structure'). Topics Maps mainly maps reality, with the possibility of including more specifically a map of documentation. The topics are *in* the map. A topic is the building block for mapping. You need more than one topic for a realistic map, hence the plural.

It's somewhat like — talking about — painting in oil. 'Oil painting' is the technique. Or, more fundamental even, oil paint is the basic material applied for painting. Likewise, pencil drawing is drawing with a pencil. It's not supposed to mean a drawing of a pencil. Or is it? Of course, context.

I'm sure that many people will experience the same problem that I had. "Topic Maps" sounds fine. It offers the solid promise of radically new possibilities. What it really is, is another label for thesaurus

with added flexibility. As I believe the thesaurus to be a major invention for information management, labeling something as a thesaurus tool is not at all demeaning. But I do want to be realistic.

I want to comment further on the claim made for Topic Maps that it supports direct modeling. I completely agree. As figure 3 illustrates, the orientation at individual documents does no longer control mapping, at least not with priority. For all catalogues, and certainly from a thesaurus on, reality is mapped directly. With improved expressive power, theoretically even at the limit — see the next section about the two basic building blocks which can also be recognized at the fundamental level for Topic Maps — of what can be imagined, Topic Maps is even less restricted by its roots of archive and library. So, yes, Topic Maps qualifies for direct modeling of reality. But witness development in information management spurred by digital technology, countless other modeling paradigms have preceded Topic Maps.

Welcome, too, to the community of conceptual modeling. How does Topic Maps compare outside its original community? Can it match, or even surpass, its competition? In its new all-encompassing environment it will need to jettison some counterproductive heritage. Does what remains still offer differential advantages?

#### **4. Intermediary remarks**

---

A problem in evaluation often is that, say, levels can get mixed up. This risk is certainly present with Topic Maps.

At its fundamental level, Topic Maps argues that mapping only requires two building blocks. They are the topic and the association. An association is upgradeable to a topic, too, and so on. Topic Maps therefore shares its core, or fundamental level, with graph theory. Topic Graphs? A topic is equivalent to a node, or vertex. An association or relationship is equivalent to an arc, or edge.[5]

But Topic Maps is of course more than its two basic building blocks. It also supplies rules and recommendations on how to construct proper maps with them. Then, it is primarily the requirements assumed at this assembly level that have subsequently by abstraction led to — (re)invention of — the set of necessary and sufficient constructs.

The assembly level is where the actual maps are developed, maintained, etc. Those are maps of reality. It simply follows that Topic Maps' rules & recommendations for assembly reflect an ontology. What is the structural build-up of reality for adequate maps to facilitate 'navigation'?

On the ontology of Topic Maps I have several critical remarks. It is important to understand that criticism at assembly level doesn't necessarily extend to the fundamental level. In fact, I already hope to have made abundantly clear that I have no argument at all with Topic Maps as far as its basic building blocks are concerned. Metapattern as a branch of graph theory, too, naturally 'goes back' to the same two elements.

This distinction between fundamental and assembly level implies that any criticism should first of all be answered at the level it pertains to. It's precisely what I am doing here, and I can only try to be as consistent as possible. But it also means that any ontological alternative I propose shouldn't meet with the reply that whatever alternative doesn't really matter, anyway. It wouldn't, the argument may go, because the basic building blocks of Topic Maps can deal with every known ontology. I already know they can, at least as long as representation through graph theory is possible. What I am making, though, is a point at the level of assembling basic building blocks, i.e. for constructing maps.

Answering criticism from the fundamental level — again, which Topic Maps shares with numerous other theories — to criticism at the assembly doesn't satisfy. Then, the issue is merely avoided. What distinguishes Topic Maps is a particular ontological position; it implies specific directions for mapping. In the next section, I will comment upon Topic Maps' directions for assembly of building blocks into actual maps, that is, upon its rules & recommendations, *not* upon the building blocks themselves.

## **5. Heritage of names**

---

Not surprisingly given its roots in language use consolidated in documents, Topic Maps maintains a formal fixation on names. If I were granted only three or four more sentences, and each one not too long, to round up this paper, I would say:

1. All right, most assembly guidelines aim to resolve naming ambiguity for actual maps.
2. So, Topic Maps acknowledges the problem but fails the accurate solution.
3. As long as the fundamental fixation is not remedied, however, maps in the Topic Maps' way are at best overly complex.
4. True enough, the building block of topic doesn't need to be filled with a name.
5. When looking at the assembly level, though, Topic Maps formally prescribes a heavy use of names and bringing up the wider potential from the basic level is therefore no proper argument for Topic Maps as it stands.

I am glad to be able to continue. Even smuggling in an extra sentence isn't enough to explain where and why Topic Maps goes wrong with names.

To get started on explaining how Topic Maps deals with information structure, I return to what I originally found confusing about its overall label; also see § 3, above. According to an enneadic framework for semiotics,[16] I now understand that topics and their associations are units applied along the dimension of signs (also read: model, map, etc.) for representing subjects (nowadays popularly read: objects). As in map *of* Europe, associated topics are assembled into a map *of* subjects. The label Topic Maps therefore effectively reflects the perspective of a map designer, builder, and so on, while the label Subject Maps (or nowadays, then: Object Maps) would bring out the view of a map

*user*. Yes, of course I do believe names are important. Substituting Subject Maps for Topics Maps helps me finding ground for explaining ‘what it does.’

A subject does not reside in a map. It is mapped. Mapping is achieved through topics and associations. However, this is right away where a credible comparison to a geographical map ends. Such a map goes a long way with a one-to-one correspondence; one geographical element, a particular ‘real’ town for example, is represented by a single typographical element, in this case probably a dot of some size. A subject map based on Topic Maps is quite different. One subject may still be represented by one topic, but that would amount to a straightforward index, only. The idea of Topic Maps that takes it beyond an index is that a veritable collection of associated topics may be involved in describing a particular subject. It is really that one particular topic is established as a kind of *center* for representing the subject. According to the tradition from which Topic Maps developed, basically a name occupies this center. A name is therefore considered the primary subject label. Rather than calling it a subject name, the International Standard refers to topic name, or *topname* for short. From several types of topic names, the so-called base name is the one placed at the center for subject representation. (I’m well aware that I’m cutting some corners here, but I return to unambiguous identification of the representational center later on.)

Then, (other) subject characteristics are added. It should come as no surprise by now that the International Standard for Topic Maps doesn’t refer to subject characteristics but to ... topic characteristics. Even a small dose of semiotic discipline would go a long way making the Standard easier to understand.

And speaking about international standards, it escapes me why Topic Maps should be elevated to such status. My impression is that it resulted from the limited perspective of a special interest group which goes against my concept of ‘standard.’ There should at least be the obligation, I feel, for including an analysis of how the new standard stands to existing standards. Of course, with a new standard offering a better solution, earlier standards should then be adapted, or withdrawn. As it is with Topic Maps, there is no mention at all of, leave alone a proper analysis of its relation to, related international standards, i.e. for information modeling at the conceptual level. I would find it normal to expect from the International Organization for Standardization to maintain consistency between separate standards. I cannot take Topic Maps seriously *as a formal standard*, for it completely lacks explicit integration with earlier standards with the same or a related goal.

Anyway, adding characteristics, or properties, works as follows (I think). A relationship, called an association in terms of Topic Maps, connects the original center to another particular topic.

Now, an interesting question presents itself. How could the association be established? Doesn’t the very relationship presuppose existence of the other topic? But, then, how did that topic become an element of the map? It figures it is the result of creating the representational center for ... another

subject. Cutting a long story short, what Topic Maps allows is the interweaving of subject descriptions. One subject's center will also act as the topic aspect — why aspect? Please note, also an association is required to complete the topic characteristic — of the characteristic of one or more *other* subjects.

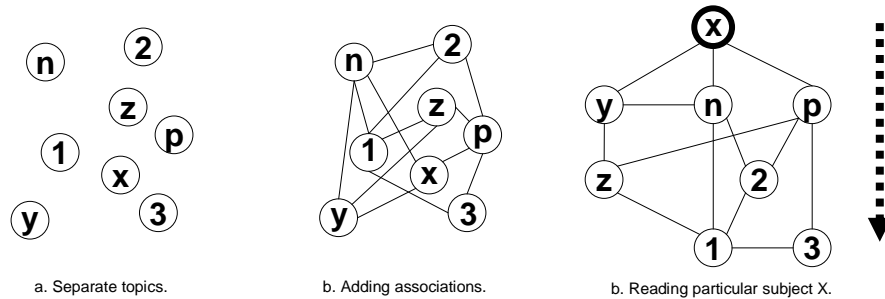


Figure 4: Reciprocal meaning through structure.

What figure 4 summarizes is, by the way, not at all a novel idea. F. de Saussure (1857-1913) developed such a concept of value for language, thereby founding structuralism.[3] I'd like to add that I find support of structuralism a strong asset. For Metapattern I even deliberately designed it.

Taking stock of where I have arrived with my analysis, a subject map constructed with Topic Maps is a structured collection of names. A particular name acts as the representational center for a particular subject. Following the relationships connecting such a chosen center to other names provides a full(er) representation of the subject in question. With *every* name a center for some subject, the number of subjects treated in a subject map equals the number of separate topics annex base names.

I'm aware that this must be not quite right. It escapes me, though, how subject, topic, base name and scope — see the next section for scope — are numerically related. Does topic as a subject's representational center allow for several base names per subject, i.e. one base name for each scope? In such a case of multiple base names, the so-called subject descriptor is required. Then, is the subject descriptor akin to a privileged base name, that is, the one with so-called unconstrained scope? More on unconstrained scope further on, too.

## 6. After Babel: scope

---

*The Bible* portrays Adam issuing names to all living creatures. Then for quite some time there doesn't seem to have been a problem with names, at least nothing worth mentioning. On the contrary, people could undertake building a huge tower, precisely because everything could be made clear to everyone. It looked that the tower might turn out too high for their god's comfort, though. The god confused man's language, with the intended result that coordinated action at the scale required for continuing work on the tower became practically impossible.[11]

Topic Maps originated after Babel, i.e. in a world where ambiguity rules. Yet, it is recognized that we do have to coordinate some activities despite linguistic ambiguity. How does Topic Maps cope?

I'll stick to considering names. As it might happen, for any number of different subjects a claim could be made for the same name as the representational center. In fact, every claim is most likely awarded. Aristotle (384-322 BCE) already points this out in the opening sentences of *Organon*.<sup>[2]</sup>

Then, with equal base names, how can one center be said to refer to one subject, and another center to another subject? Topic Maps allows an explicit scope to be added to a center. When two names are the same but 'carry' different scopes, they are taken to refer to different subjects.

Or is it, that one center may 'contain' several base names? And is it that only base name and scope maintain a one-to-one correspondence? It doesn't make any difference for my argument, as it centers around names, anyway. A scope equals a so-called namespace.

A scope for a (base) name is also constructed by association(s) to (other) topics. When constituting scope, Topic Maps refers to such topics as themes. I don't believe such an extra label helps keeping a clear view of scope, so I won't use it. As I indicated above by mentioning De Saussure's structuralism, a theme — or what I would prefer to call a scope topic — involves a referential center for yet another subject.

I repeat my admission to an important gap in my theoretical understanding of the International Standard for Topic Maps. Is it possible to just create several topics and all fit them with the same name? If so, is there some 'internal' system for identifying topics, with whatever name(s) added on merely an external property that may be changed at will?

It certainly is a formal constraint set by Topic Maps that different subjects cannot be represented by the same (base) names when those names' scopes completely match. When this condition of name/scope equality is met, the assumption of different subject is considered false. As a remedy, topic characteristics are merged, resulting in representation of a single subject.

This is the point, I find, where an imbalance of Topic Maps manifests itself most clearly. Let me first state what I believe that the designers of Topic Maps assumed about (dis)ambiguity. It certainly seems that a name is supposed to be unique as a matter of principle. Of course I might be wrong about such an assumption, but I'll continue as though I'm right. I find it corroborated by Topic Maps' concept of unconstrained scope. It is the default scope and doesn't require any scope topic (theme) to be specified. I would say such an arrangement for default only makes sense when base names normally differ.

Combined with the property that scope is the set theoretical union of scope topics, two discontinuities occur in Topic Maps' approach to scope. For without any scope topic specified, scope is considered universal. Then, on the basis of a single scope topic, scope makes a jump across the specter for range and suddenly becomes minimal, i.e. limited to that particular added topic. With every further topic

added, scope increases by union of the scope topics. Suppose that a particular scope is all of a subject map except for only one topic. Then all but that particular topic need to be specified as scope topics. For the next step, however, that is, making scope universal, suddenly all scope topics may be eliminated as no specification is the default for comprehensive scope. I favor a mechanism where more specifications for situation annex context makes behavior (see more about my emphasis on situational behavior, later on) *consistently* more specific, rather than less.

While I understand where scope's default comes from, I therefore still miss why scope results from the union of scope topics. As I said, I strongly prefer precision in scope and would therefore have favored intersection of scope topics as the rule for scope specification. Topic Maps doesn't overlook the need for intersection, but recommends an additional topic is created for that purpose. This exception for Topic Maps is the very rule for Metapattern. Making exceptions possible for what has developed into the rule at the ever increasing scale of integrated information management might be a theoretically valid response to criticism, but soon stops to be practically viable.

Putting it bluntly, the naming principle of Topic Maps suffers from myopia. Only small parts of the world (also read: reality) can be 'mapped' without running into difficulties with names or, in general, multi-valued signs. That is why Wittgenstein (1889-1951) introduced the concept of language game, which is the limited part of the world where language use is supposed to be unequivocal.[19] A so-called community of practice — a label I've already applied; see above — involves more or less the same idea.[13]

So, my reconstruction is that subject maps originally had limited coverage. Now that was, please note, *before* Topic Maps. The small scale of each mapping sort of permitted that names could be counted upon to form representational centers for subjects, just like that. In fact, a particular subject map would most likely assist a particular language game annex community of practice, only. And the historical development of whatever game/practice would already have eliminated the most obvious conflicting claims for names—or it wouldn't be a language game, now would it? Topic Maps came after, because it really doesn't answer a need for better subject maps for separate language games.

A need for additional control mechanisms arises when previously separate subject maps are merged. I have the distinct impression Topic Maps was initially *primarily* designed to meet the requirement of joining smaller maps into a single larger map. But I suppose those still were maps of the same territory, i.e. where convergence of meanings for single names was still practical. The benefits are clear from the perspective of gaining access to relevant documentation. After merging, only a single map needs to be consulted for navigating.

At the larger scale of synthesis, though, language games effectively tend to get mixed up. As a consequence, a name no longer practically guarantee a one-to-one correspondence to a subject. My reconstruction of the design of Topic Maps leads me to think that the naming principle was left intact,

but some additional measures were specified for, say, original subject maps. What Topic Maps essentially provides is a prescription for separate subject maps, i.e. adding constraints, in order for their merger into a comprehensive subject map to succeed. What counts as success is that no name becomes ambiguous.

I repeat that Topics Maps, seen from this perspective, is first of all about preventing ambiguity when subject maps are merged. Topic Maps governs that conditions for prevention should already be secured within each and every original subject map. With the proper conditions fulfilled beforehand, the merging process can proceed automatically (which I recognize as a primary requirement for Topic Maps).

Why do I deal at length with the origins of Topic Maps? I want to bring out what has developed into a contrast. Today, Topic Maps is increasingly promoted as a method for information management per se.[8, 9] I mean that the original requirement for accurately merging subject maps for at the most slightly disjunct territories (also read: subject areas) runs the risk of disappearing from view. While the purpose of merger is no longer dominant, rules & recommendations remain unchanged. But as the ambitions for Topic Maps shift, does it still offer an optimal solution?

## 7. Different perspectives on behavior

---

My question ending the previous section is of course rhetorical. My most general criticism is that Topic Maps builds from a traditional core that is only suitable for — the illusion of — straightforward mapping, that is, at a scale where ambiguity is still absent (give or take an exception for which some additional measures still don't overburden the basic structure). In the face of new opportunities, Topic Maps tries to expand a solution for a simple problem into something for solving a more complex solution. My hypothesis reads that merging simple maps is the problem that Topic Maps originally set out to solve. It accomplishes it through measures for additional structure in maps before they are being — automatically — merged. But it still is all *around* names.

As I argued, Topic Maps takes a (base) name as the representational center for a particular subject (see figure 4). Among one and the same name's characteristics several so-called scopes may appear. This amounts to scope being a property of name. It exemplifies what I've been trying to express as my criticism. The *effect* of scope is to direct name to a correspondence to a particular subject. In terms of behavior, Topic Maps thus actually prescribes that name follows scope. However, its structure suggests otherwise, with scope added to name.

At least my first impression is, when scope is made into a property of name, that scope also follows name in behavior. No, it does not, and I completely agree. But then I disagree with how the rules & recommendations of Topic Maps decree structure. Its heritage of names forms an obstacle to variety control at the scale and pace of the, let me say, information society. Practical tools for complex



information management can no longer result from merging simple, isolated ‘maps.’ They have to be designed from the ground up because *structural* ambiguity requires measures of a qualitatively richer order of variety. (Yes, isolated maps may be incorporated into the encompassing information infrastructure, allowing for gradual and therefore manageable change.)

In the information society of now global proportions, ambiguity is not an unpleasant exception that can be somewhat circumvented. It turns from an exception in countless separate tools into the rule for a genuine information infrastructure. Indeed, language use is basically ambiguous. It therefore must be treated as a rule for our tools for information management. Shortly, I will indicate roughly how Metapattern handles ambiguity. Let me first finish my analysis of Topic Maps.

Another clear indication of how Topic Maps represents a perspective common to the community of archivists and librarians is its meaning of type. The IT community mainly considers type a mechanism for efficiency when prescribing behavior. The rough idea is that several ‘items’ may exhibit the same behavior. It would then be redundant to include instructions for every individual item. An abstraction is therefore added, called class or type. Instructions are subsequently specified for the type, which a particular item consults for ‘its’ behavior. This way, an item is a member of a type. What I want to emphasize is that type holds a superordinate position relative to item; efficiency of behavioral specification is the organizing principle.

The A&L community traditionally uses type or category for improving relevance in selection. Please note that there is nothing attached to type, such as a specification of behavior for all pertinent items. It is therefore not more efficient to superordinate type over item. The choice is basically neutral, but there are ontological consequences (which is my point). In practice, type appears as subordinate to item, that is, primarily specified as an item’s property. This way, please note, an unlimited number of types may be specified for any item. It results in that item being selected whenever a search is made for any of its types. So, multiplicity of type is advantageous for relevance in retrieval.

Multiplicity should be eliminated for selection of behavior, however. For an item’s machine-generated behavior must be determined without ambiguity (which explains why the IT community ‘positions’ type differently).

Topic Maps maintains the perspective of the A&L community. So, topic type is yet another characteristic/property of topic. As could be expected, it is also established through an association with another topic (necessarily acting as the representational center for yet another topic, and so on).

My exposition on how Topic Maps applies typing should once again serve to stimulate an awareness of which classes of applications it is best suited for, and which not. With both subordinate and thus — the possibility of — multiple types, type in Topic Maps does not qualify as the mechanism for controlling — performance of — behavior. A subject map remains predominantly static. It represents subjects, more than that it acts, too. Topics are not agents.

## 8. The case for requisite variety

---

Metapattern was developed independently from Topic Maps, but what I did bring to it was a deep appreciation of long-standing methods for managing archives and libraries.[14] I added some maxims derived from, in my case anyway, applied mathematics.

First and foremost is that real variety deserves utmost respect. It simply follows that whatever tool must exhibit requisite variety, meaning that it matches the complexity of that part of the world where it is destined to perform. Secondly, and contrary to popular belief, requisite variety is impossible to develop stepwise, i.e. from simple beginnings and then just adding on properties until it matches the real variety. It should always be recognized that for any real complex problem, variety is *sui generis*. Then, synthesis is not the result of incremental steps. At least the design needs integrative quality. Of course, a design may fall short. But whatever next design always needs to address requisite variety as a demand for integration.

Many practical problems are of course both too large-scale and too complex for detailed design right from the start. Then, a deliberate attempt must be made that at least the foundation is supplied with such variety as to minimize the risk that it inadvertently should be broken down to make room for what next promises to carry design etcetera to eventual completion.

Above, I've already stated that it is difficult to go wrong with just node and relationship as necessary and sufficient basic building blocks for modeling. To be sure, Metapattern builds from these, too.

What gives Metapattern its distinct, say, flavor rests with the specific rules & recommendations I've designed for it. My lengthy analysis of Topic Maps now allows me a very short statement about Metapattern's method for information modeling.

Metapattern views *behavior* as the choice about which ambiguity cannot be allowed. A particular behavior must therefore be determined *before* it can be called upon. Ambiguity exists where a particular object/subject may exhibit different behaviors. Then, what counts as the arbiter for choosing a particular behavior from the object's behavioral repertoire? Metapattern holds that a particular-object-in-particular-situation directs a particular behavior. This is exactly the opposite of structure for Topic Maps. With Metapattern, behavior is a property of object-in-situation. It follows that situation constitutes a behavioral type, at least when all relevant objects behave similarly in that situation.

For Topic Maps, scope is still optional. There, it is 'only' to be included for combating ambiguity. But situation is not at all optional with Metapattern. Neither is what Topic Maps considers a subject descriptor. *Every* behavior is situational. *Every* object-in-situation, that is, the combination of an object and a situation, is uniquely identified (which allows structural dynamics in time, a treatment of which would really extend the scope of this paper; see especially [16] and [18]). Nothing hinges on external names; when required, name is another property of an object-in-situation.

The very definition of situational behavior implies that all ambiguity is eliminated as a matter of Metapattern principle. Metapattern's structure is fundamentally 'open.' It supports information models, say maps, at whatever scale. What can always be added, or modified, are first of all disambiguated situations. Everything else follows.

Once a situation is modeled, and what corresponds to situation is context, an object's model for situational behavior is unequivocal. From an afterthought with scope, ambiguity control becomes forethought with situation. Pepper seems aware of shortcomings, but doesn't know where to look for a solution. He mentions "a gut intuition that scope as a flat set of themes is insufficiently expressive. [...]he expression of scope as a flat set of themes is less than satisfactory. [...] It seems that scope may need *structure* in order to be truly useful." Quite right! And Pepper is mistaken "that there is as yet no well-defined way of structuring scope." [8] Whereas Topic Maps informally distinguishes between scope and context, modeling with Metapattern takes only context for sorting out a specific behavior. There are of course consequences for how context etcetera is, indeed, explicitly structured as a map of situation.

A situation is also modeled with nodes and relationships. And so is behavior. Metapattern actually treats situation, object and behavior as relative concepts. Somewhat reminiscent of Ranganathan, recursion of node/relationship keeps the overall structure compact. So, what Topic Maps has in common with Metapattern is consistent application of just two basic building blocks. With rules & recommendations explicitly aiming at controlling larger variety, Metapattern incorporates the opportunities that Topic Maps affords. Or, at the assembly level Topic Maps is a logical subset of Metapattern. The other way around, Topic Maps would need to abandon its most characteristic rules & recommendations. Resorting to manipulating its basis building blocks only, for which Metapattern offers shortcuts, the difference would then come out as a great loss in efficiency. And the question would become appropriate whether, by recourse to generally known basic building block, only, is it still really is Topic Maps that is applied.

It lies beyond the scope of this paper for me to provide any serious introduction to Metapattern itself. What remarks I made in this final section mainly serve to support my argument that for increasingly complex issues in information management Topic Maps does not qualify. New opportunities require tools with matching sophistication: Metapattern.

## Literature

---

1. Aitchison, J., and A. Gilchrist, *Thesaurus construction: A practical manual*, 1972, Aslib, second edition 1987, reprint 1990.
2. Aristotle, *The Works of Aristotle*, Encyclopaedia Britannica, two volumes, 1952, 29<sup>th</sup> printing 1987.
3. De Saussure, F., *Course in General Linguistics*, originally published in French 1916, Open Court, 1989.
4. International Organization for Standardization / International Electrotechnical Commission, ISO/IEC 13250 Topic Maps, ISO/IEC, second edition 2002.
5. Johnsonbaugh, R., *Discrete Mathematics*, Macmillan, 1894.
6. Park, J. (editor), *XML Topic Maps: Creating and Using Topic Maps for the Web*, Addison-Wesley, 2002.
7. Peirce, C.S., Logic as semiotic: the theory of signs, 1897-1910, in: *Philosophical writings of Peirce* (editor J. Buchler), 1955, pp 98-119.
8. Pepper, S., Towards a General Theory of Scope, Ontopia, 2001, available through [www.ontopia.net](http://www.ontopia.net).
9. Pepper, S., The TAO of Topic Maps: Finding the Way in the Age of Infoglut, Ontopia, 2002, available through [www.ontopia.net](http://www.ontopia.net).
10. Ranganathan, S.R., *Colon Classification*, Sarada Ranganathan Endowment for Library Science, 6<sup>th</sup> edition, 24<sup>th</sup> reprint 1998.
11. *The Bible*, London, UK, Hodder & Stoughton, New International Version, edition 1989.
12. Wellisch, H.H., *Indexing from A to Z*, H.W. Wilson, 1991.
13. Wenger, E., *Communities of Practice: Learning, Meaning, and Identity*, Cambridge University Press, 1998.
14. Wisse, P.E., Multicontextual paradigm for object orientation: a development of information modeling towards fifth behavioral form, originally written in Dutch 1991, English translation in: *Informatiekundige ontwerpleer*, Ten Hagen & Stam, 1999, also available through [www.wisse.cc](http://www.wisse.cc).
15. Wisse, P.E., *Metapattern: context and time in information models*, Addison-Wesley, 2001.
16. Wisse, P.E., *Semiosis & Sign Exchange: design for a subjective situationism, including conceptual grounds of business information modeling*, Information Dynamics, 2002, also available through [www.wisse.cc](http://www.wisse.cc).
17. Wisse, P.E., [Information metatheory](#), in: *PrimaVera*, working paper 2003-12, Amsterdam University, 2003.
18. Wisse, P.E., [The pattern of metapattern: ontological formalization of context and time for open interconnection](#), in: *PrimaVera*, working paper 2004-01, Amsterdam University, 2004.
19. Wittgenstein, L., *Philosophical Investigations*, MacMillan, two-language edition (English-German), 1953.