Comparison of multiple taxonomic hierarchies using TaxoNote

David R. Morse[1]
The Open University,
United Kingdom

Nozomi Ytow [2]
University of Tsukuba,
Japan

David McL. Roberts[3]
The Natural History
Museum, United Kingdom

Akira Sato[4]
University of Tsukuba,
Japan

Abstract

Recent work on modelling taxonomic names and the relationships between them has highlighted the need for capturing the multiple names and hierarchies that exist in taxonomic nomenclature. In this paper we describe TaxoNote Comparator, a tool for visualising and comparing multiple classification hierarchies. In order to align the hierarchies, the Comparator creates an integrated hierarchy containing all the taxa in the hierarchies to be compared, so that alignment of the hierarchies can be maintained. In addition, a table of assignments reports the taxonomic names that are common to all hierarchies and the differences between them, which facilitates structural comparisons between the hierarchies.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques – Graphics data structures and data types; J.3 [Life and medical sciences]: Biology and genetics

Keywords: taxonomy, nomenclature, visualisation, rough set theory, formal concept analysis.

1.   Introduction

The rules of taxonomy are quite specific and are well documented in the taxonomic codes of nomenclature (e.g., for animals, the International Code of Zoological Nomenclature [Ride et al. 1999]). Taxonomic understanding, however, is dynamic, leading to constant changes in the taxonomic entities, the taxonomic names applied to those entities, and the relationships between the entities. Consequently, taxonomic concepts and the application of names to those concepts varies through time.

Recent work on modelling taxonomic names and the relationships between them has highlighted the need for capturing the multiple names and hierarchies that exist in taxonomic nomenclature. A number of projects have considered this problem, including Hiclas [Jung et al. 1995], Nomencurator [Ytow et al. 2001], Prometheus [Pullan et al. 2000] and IOPI [Berendsohn 1997]. Data models that incorporate multiple hierarchies will be crucial in facilitating the effective integration of biodiversity data from diverse sources, since multiple and overlapping taxonomic concepts must be tracked, as well as the names that have been applied to these concepts. Equally important are visualisations which permit the comparison and exploration of several hierarchies simultaneously.

[1]e-mail: d.r.morse@open.ac.uk [2]e-mail: nozomi@biol.tsukuba.ac.jp
[3]e-mail: dmr@nhm.ac.uk [4]e-mail: akira@cc.tsukuba.ac.jp

1.1.   Definition of Terms

Classification hierarchies are a means of organising information by groups. The rules of nomenclature insist that within the scope of the rules, broadly animals, plants and prokaryotes, names must be unique. The internal nodes of such hierarchies represent successively more general groups. Phylogenetic hierarchies, commonly referred to as trees, are statements of an evolutionary hypothesis and ideally should be constructed from bifurcating nodes. All living taxa are represented as terminal nodes on the tree, internal nodes represent ancestral taxa. Higher taxonomic names are given to sub-trees called 'clades' rather than specific nodes. In keeping with the tree analogy, terminal taxa are sometimes referred to as 'leaves', small groups of taxa as 'twigs' and large groups as 'branches'. Modern systematics seeks to represent evolutionary relatedness within an hierarchical classification.

1.2.   The Taxon Concept

The objects to be classified begin with individual specimens, which are grouped into species. Systematics is fundamentally an ostensive process, i.e. classification by example. For instance, an object is considered to be a chair because it looks like objects that we know to be chairs. As an aside, works of art often explore this conceptual boundary. Dictionaries, on the other hand, seek to define objects with a summary of their properties, an attributive definition, which delineates a boundary within which an object must fall to warrant the use of the name. If we imagine this process as casting individual objects into an attribute space, then ostensive definitions define points within a cluster and lend their name to the cluster, whereas attributive definitions define the boundary of the cluster. The number of clusters recognised in a given volume of attributive space reflects how well populated that space is. If there are few examples, it is difficult to define clusters at all. If there are many examples, then clusters may be well delineated if the appropriate attributes have been measured.

When first defining taxa it is inevitable that the attribute space will be sparsely populated, so it is difficult to assess the adequacy of the choice of attributes. Consequently, many definitions are of the form "like so-and-so except …". Assignment of examples to such a definition will therefore be an expression of an author's opinion and will perforce modify the taxon concept to a greater or lesser extent.

Taxa are grouped for convenience of handling and there is a recognised seniority in the major ranks from Kingdom through Phylum, Class, Order, Family, Genus to species. There is an almost inexhaustible supply of ranks between these groups principally super-, sub-, and infra-, as well as ranks such as tribes, varieties and strains.

A name literal is a tag which is associated with the taxon concept and establishing the equivalence of names means both establishing equivalence of name literals and establishing the equivalence of the taxon concepts that they represent.

1.3. Problem Scope

Taxonomic hierarchies are a sub-set of hierarchical structures and are similar, but not identical, to familiar constructs such as file systems. Taxonomic hierarchies comprise an organised list of taxonomic names that are drawn from diverse data sources and organised according to an expert in the local domain. The scope of coverage need not be global and is often geographically restricted, the British Fauna, for instance. Publication of new taxa after the publication of the main hierarchy will often, but not necessarily, specify the hierarchical position that the author intends for the new taxon. Various experts will inevitably propose a variety of such views. In [Ytow et al. 2001] we presented a model of taxonomic nomenclature that was designed specifically to be able to manage taxonomic names that are organised into multiple hierarchies. In that paper we also described software that implemented a prototype of the model.

There is a distinction to be drawn between the types of hierarchy that are to be compared. Phylogenetic trees are formal statements of hypotheses of evolutionary relationships and as such often represent alternative arrangements of a given set of leaves on a tree, which means that any given leaf can be found somewhere on the trees being compared. This is the problem that [Munzner et al. 2003] addressed. Classification systems, on the other hand, are hierarchies which encompass potentially different leaves, since the validity of each taxon in the classification is a statement of judgement by the author building the hierarchy. This means that missing and incompatible data are much more prominent components of the latter case, which we explore in this contribution.

1.4. Implementation environment

We chose to carry out our development work in Java, for the same reason that [Munzner et al. 2003] chose Java: its support for multiple operating system platforms. While our software works with the entire classification data sets presented in the InfoVis 2003 Contest, our visualisation tool worked best if it had at least 1.5GB main memory available. Such systems are not readily available to taxonomists, so we chose to work with the more manageable mammalian subsets of the hierarchies. However, we have not attempted to optimise memory management, so we anticipate that future versions of our software will use system resources rather more efficiently than the version reported here.

1.5. The Nomencurator project

In [Ytow et al. 2001] we described a data model that supports multiple taxonomic views and Nomencurator, a prototype implementation of the data structures. Since then, we have been working on TaxoNote, which is a graphical user interface to the Nomencurator data structures. TaxoNote (short for **Taxo**nomist's **Note**book) was conceived as an extension to the familiar laboratory notebook. In addition to the core taxonomic names database, TaxoNote should provide tools that support the visualisation and comparison of multiple taxonomic hierarchies.

It is the latter function that will be described in the present paper, particularly in the context of the IEEE InfoVis 2003 Contest. We will discuss the issues involved in providing visualisations that enable taxonomists (domain experts rather than novices) to visualise and to work with such hierarchically organised data.

2. Hierarchy visualisation and comparison

In general, trees are specified by a pair consisting of a set of nodes in the tree and a set of ordered relationship between these nodes, i.e. parent-child relationships. A taxonomic hierarchy is a kind of tree in the general sense, and hence we can divide comparison of hierarchies into two tasks: comparison of nodes and comparison of relationships between the nodes. These two tasks are convolved in taxonomy because of the recursive nature of taxon concepts.

In studying a hierarchy, a taxonomist might wish to know the following information about a taxonomic name: who is my parent, what is my hierarchical position (the chain of parents back to the hierarchical root), who are my siblings, and who are my children? Additionally, in comparing hierarchies taxonomists are particularly interested in areas of conflict rather than in areas of agreement, and in the principled exploration of structural differences between the two (or more) hierarchies. As noted by [Munzner et al. 2003], current visualisation techniques for large trees do not support these tasks particularly well, since the tools and techniques used are better attuned to support browsing rather than targeted navigation of the hierarchies.
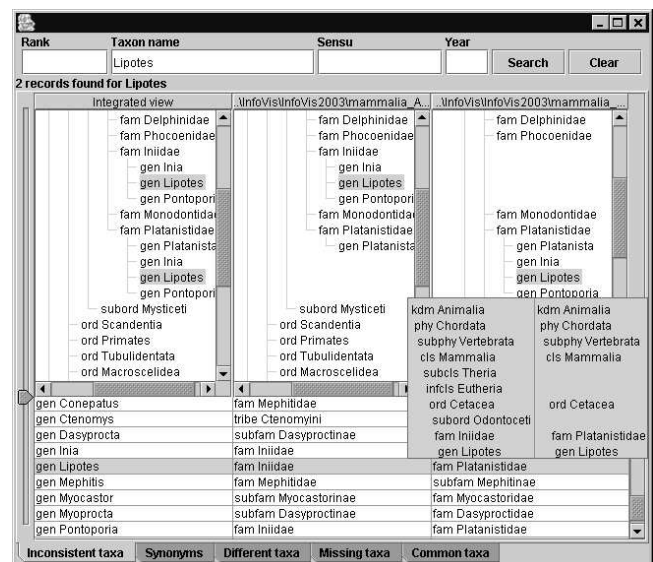

Figure 1. The hierarchy visualisation and comparison tool within TaxoNote.

A screen dump of the TaxoNote hierarchy visualisation and comparison tool, the Comparator, is shown in Figure 1. The Comparator display can be divided into three parts:

- A **Query** panel at the top can be used to search the hierarchies that are being displayed for particular taxonomic names, by text entry.

- A **Hierarchy Comparison** panel shows the two hierarchies that are being compared (centre and right) and an 'integrated view' (left) where the hierarchies have been merged into one, composite, hierarchy. An additional pane would be added for each hierarchy being compared by the application. The hierarchy comparison panel provides a list of siblings and children of a taxon. It also captures the parent taxon and the path to the hierarchical root. These may not be displayed if there are many siblings or children of a node, (e.g. the genus *Eucaryptus* which has more than 512 species as child taxa).

A **Pop-up panel** gives a short summary of the path to the root in these cases (Figure 1).

- An **Assignment Table** at the bottom shows various alternative views of where names that appear in the hierarchies are assigned. It contains information on the parent taxon and potential equivalence of taxon concepts depending on its modes. While the **Hierarchy Comparison** panel gives a top-down oriented view, the **Assignment Table** gives a bottom-up oriented view.

2.1. The Query Panel

In any large data set, searching for a particular datum by eye is tedious, so efficient mechanisms such as search tools are necessary to focus the display and the user's attention on the area of interest. As the Comparator is particularly designed to be a tool for taxonomists, rather than a general tool for browsing trees, so additional fields to the taxon Name are included as potential query fields. These fields would normally be used in addition to the taxon name in order to refine further the search. The other search fields that are available are the taxonomic Rank, Sensu and Year. The latter two fields are metadata which are important in modelling multiple taxonomic hierarchies, since they allow you to compare, distinguish between and reconcile different taxonomic opinions of the taxon concepts that are linked to the same taxonomic name [Berendsohn 1995]. It should be noted that such additional metadata were not present in the InfoVis data sets, which had implications for the nature and complexity of the comparison algorithms that were used when the two hierarchies were compared. We will return to this point below. Finally, it should be noted that the inclusion of these other search fields, including taxonomic rank, may not make sense in other hierarchy visualisation applications such as file system hierarchies. In these situations, alternative mechanisms may be required in order to focus the search.

2.2. The Hierarchy Comparison Panel

In Figure 1, notice that we have prefixed all names with an abbreviated form of the taxonomic rank as an aid to navigation and comparison. This makes sense in taxonomic hierarchies where different levels in the hierarchy have different names (ranks), but in the majority of other hierarchies, intermediate levels between the root and leaves of the hierarchy are not distinguished. Also note that as with the Microsoft Explorer interface, additional levels of the hierarchies can be expanded and contracted at will.

We chose an indented representation for the hierarchies in the Hierarchy Comparison panel because this is extremely familiar to taxonomists. Hierarchies have been presented this way in taxonomic publications for hundreds of years, so much valuable source data are available in this form. This representation is familiar to most computer users through applications such as Microsoft Explorer. While other representations such as Hyperbolic Trees and TreeMaps [Bederson et al. 2002; Graham and Kennedy 2001] may have a higher information density, it is important that the names retain their visibility and readability at all times.

A key benefit of the indented method of hierarchy display is the ease with which two or more hierarchies can be compared visually by arranging them side by side in columns. We did consider arranging the two hierarchies as mirror images of each other, but rejected this proposal because of the potential difficulty in interpreting the mirrored hierarchy, and the inapplicability of this approach to the comparison of more than two hierarchies. In the reverse hierarchy, more deeply nested levels in the hierarchy appear further to the left, contrary to the way hierarchies are usually represented in Western cultures. We note in passing that in languages which are written from right to left (e.g. Arabic), it would be more natural for all hierarchies to be shown in this way.

2.2.1. Alignment of taxonomic names

Core to the problem of alignment is establishing the Best Corresponding Node (BCN, see [Munzner et al. 2003]). Ideally, corresponding nodes would represent equivalent taxonomic concepts. Unfortunately the taxonomic concept itself is extremely difficult to pin down [Ytow et al. 2001] and is approximated in one of two ways, either by consideration of the objects (taxa or specimens) included in the concept [Munzner et al. 2003; Pullan et al. 2000] or by analysis of the attributes of the taxon, i.e. the shared characters of the group. The former method is very sensitive to the contained set being incomplete for any reason, and data for the latter method are rarely available. Other proxy measures of the taxon concept have to be combined to establish the BCN, which include the hierarchical position (parent list), the included objects (the child list), but interpreted in a flexible manner, where positive matching counts for more than missing data and absence of conflict counts in favour, conflict against. This set of relationships is subtle and is currently being explored using rough set approximations and formal concept analysis [Yao et al. 1997; Ganter and Wille 1999].

In the Hierarchy Comparison panel, rows which are aligned have the same names in the same hierarchical position in both hierarchies (e.g. family Phocoenidae in Figure 1). Rows which are not aligned are indicative of names missing from one hierarchy, perhaps because they are newly created (e.g. family Iniidae) or names whose hierarchical position has changed from one hierarchy to the other (e.g. genus *Lipotes*).

2.2.2. The Integrated Hierarchy

In order to align the two hierarchies and to maintain their alignment while the display panels are scrolled, a consensus hierarchy is constructed from the source hierarchies that are being compared. In areas where consensus is impossible, the name literals are duplicated and entered in each potential position (e.g. genus *Lipotes* in Figure 1). This process requires the establishment of the BCN for each taxon in the integrated view.

We noted earlier that a taxonomic name is composed of its name literal and the taxon concept. Therefore, we need to establish the equivalence of both the name literal and the taxon concept. Evaluation of the equivalence of name literals is a non-trivial task. The ostensive nature of the taxon concept also makes examination of taxon concept equivalence complex because concept comparison by ostensive sets is too sensitive to the addition of a new specimen that has been identified as a member of the taxon. We used a rough set approximation to evaluate ostensive concept equivalence because it is robust to the addition or removal of a new member and also robust to insertion or removal of intermediate taxa. However, the rough set approximation is rather expensive because it requires that the whole hierarchy is searched. Hence we started from equivalence of name literals as a first approximation to the equivalence of taxon concepts and only used rough set approximations when literal matching is insufficient.

Integration of hierarchies proceeds hierarchy by hierarchy. First, we reproduce all names in the first hierarchy into a subtree under an unnamed root node in the integrated hierarchy. Integration of the second hierarchy proceeds as follows. Names in the new hierarchy are examined so that they can be merged with names that already exist in the integrated hierarchy. This is done in a top-down way, i.e. starting from the root name. If the integrated hierarchy does not contain a taxon with the same name literal and the same rank as the name under examination, then the taxon is integrated into the hierarchy under the unnamed root as a new taxon. If the taxon under examination is the root taxon and there is only one taxon with the same literal name and rank in the integrated hierarchy, then these two nodes are integrated into one taxon, and then its children are examined recursively. If not, then taxa with the same literal and rank, and an equivalent parent are looked for in the integrated hierarchy. Note that there can be more than one taxon in the integrated tree with the same literal and rank, but with different parent literal and rank. If only one parent-equivalent taxon is found, then the taxon under examination is integrated with the taxon. If there is no parent-equivalent taxon but there is a literal and rank equivalent taxon, these candidates are screened by a rough set compatibility test. If no taxon passes this test, then the taxon under examination is inserted as a new taxon into the integrated hierarchy. Otherwise, if only one taxon passed, then the taxon under examination is integrated with the taxon that passed. If two ore more taxa passed the test, then the parent paths of these taxa are examined. If a path-embeddable taxon is found, then the taxon under examination is integrated with the taxon. If this is not the case, i.e. no path-compatible taxon is found, then the taxon is inserted as a new taxon into the integrated hierarchy. If, the taxon has child taxa, then the same procedure is applied recursively. This recursive application may require subdivision of a taxon, or re-ordering taxa, depending upon the integrated structure.

The resulting integrated hierarchy is shown in the left hand panel in Figure 1, as the Integrated View. This is created in order that it can be used for node geometry calculations while the two hierarchies are being aligned, and in order to maintain alignment while they are being scrolled, expanded or collapsed. Hierarchies proposed by different authorities (taxonomists) are likely to embrace different taxonomic concepts that may or may not be identified by the same name. Therefore, establishing node equivalence is not trivial and we are still working on algorithms for constructing the composite hierarchy that is shown in the Integrated View.

2.2.3. Comparison of taxonomic names

Another research issue lies in finding effective ways of highlighting discrepancies and mis-alignment between the two trees. The genus *Lipotes* in Figure 1 shows two ways in which this can be done. In the two hierarchies, the presence of gaps in the hierarchies indicates taxa that have been inserted, deleted or moved to another taxon. In the Integrated View, it can be seen in Figure 1 that the genus *Lipotes* has been replicated in order to create the Integrated View. The necessary inclusion of duplicates of a name has the potential to be a way of indicating regions of difference between trees. Indeed, an estimate of the number of incompatible views can be obtained by simply counting the number of duplicate names in the Integrated view.

2.2.4. Scrolling

Targeted navigation, by expanding hierarchical levels and scrolling through them when looking for a given taxon in a large hierarchy, is very difficult. In contrast, browsing the hierarchy is reasonably well supported by such simple user interaction components. We implemented the query mechanism and pop-ups (described above) in order to support targeted navigation. In addition, the hierarchies and integrated view can be scrolled in concert by holding down the middle mouse button while any of the hierarchy display panes is scrolled. This facilitates the search for a particular taxon and the structural comparison of the different hierarchies.

2.2.5. Path visibility

The conventional tree display used in the Hierarchy Comparison panel by indented text has two roles: it shows the path to the root node, and the child and sibling nodes. Blank lines that are inserted for hierarchy alignment sometime make it difficult to manage these two roles in a restricted size display. This is also the case if the hierarchies are wide or deep. We used the Pop-up Panel to ensure path visibility in these cases. When the user places the mouse cursor on a name in one of the source hierarchies, a panel will pop up after a short delay containing the path information of each hierarchy shown in the aligned way. If corresponding names do not exist in one of the hierarchies, then "(no match)" will be displayed instead of the path. When the user puts the mouse cursor on a name in the integrated view, the Pop-up Panel contains only the path of the integrated node, because there can be two or more nodes having the same name, which could be too complicated to show in a rather small pop-up panel. We did consider including sibling information in the pop-up panel, but we rejected this because it duplicates information. This reflects the fact that position of a node in a hierarchy is determined by both the path to a node and the siblings of the node.

2.3.   The Assignment Table

The bottom panel contains the Assignment table which consists of a number of organised lists whose purpose is to allow the user to explore the differences and commonalities between taxon concepts in the hierarchies. The table is structured into columns, one for each hierarchy pane. The primary taxon is given on the left, underneath the integrated view while the parent taxon is listed underneath the appropriate hierarchical pane. The Assignment Table panel contains multiple tables with tabs which can be selected depending on the category of taxa which are gathered in each table.

As mentioned above, a name is a pair of a name literal and a concept accompanying the name. Two names are equivalent if and only if the name literals and the concepts are equivalent. A factorial combination of these equivalencies gives four cases, i.e. both literals and concepts are equivalent, either of the literals or concepts are equivalent but not both, or neither the literals nor the concepts are equivalent. The first and last cases implies that each unique concept has its own unique name literal, and if name literals are different then these names designate different concepts. The remaining two cases are known as homonyms if only the literals are equivalent, or synonyms if only the concepts are equivalent. These cases, which are of particular interest to taxon-omists are listed in tables with the tabs 'Inconsistent taxa' for homonyms and 'Synonyms' for synonyms. The former tab title is

used instead of "Homonyms" because taxonomists use the word homonym in a specific technical sense.

There are two primary ways for potentially equivalent taxa to differ between hierarchies: first they may be absent in one or more hierarchies and second they may be placed within different taxonomic groups, such that their parent chain is different. Other tabs at the bottom of the Assignment Table allow the user to see those taxa which are missing from one set or the other ('Missing taxa' tab), while those taxa with different positions are summarised under the 'Different taxa' tab. Those nodes in common are listed under the 'Common taxa' tab.

A pop-up window is available on the leftmost column which gives the position of the name in the table as a pair consisting of the row number of the name and the total number of rows in the tab. The total number of rows in each tab gives a summary of the similarity and incompatibility of the hierarchies under comparison.

One application of the Assignment Table is illustrated under the 'Missing taxa' tab by the species *Acomys cineraseus* (in Mammals A) and *Acomys cinerasceus* (in Mammals B), that looks suspiciously like a spelling error either in the original publication or in the data preparation.

3. The InfoVis 2003 Contest Data Sets

It is our contention that no one tool can solve all visualisations of hierarchical data problems. We have chosen to address one particular type of data – classification hierarchies – which may be characterised as being non-quantitative data. Our approach would need significant additions in order for it to perform well at visualising hierarchically arranged quantitative data; data which is often well suited to visualisations using TreeMaps [Bederson et al. 2002]. Such additions to our system could include colour-coded glyphs or bars alongside, or in place of the text labels.

Classification hierarchies are also unusual in that the names present in the hierarchies should be unique. The appearance of the same name in different places in a hierarchy is indicative of homonymy and is of interest to taxonomists as an area that requires taxonomic revision. In contrast, file system hierarchies are replete with duplicated names. Files called 'index.html' abound in websites – the file logs_A_03-02-01.xml records 3356 occurrences of this file, for example.

In classification hierarchies, the name is just that because of the assumption that taxonomic names in a hierarchy are unique. The position of the name in the hierarchy – the rank – gives extra information about the name. In contrast, in a file system hierarchy, the name consists of the path to the file in addition to the actual file name. While components of the path may give additional information about the file, this interpretation is not as strong as the rank in taxonomy. Clearly very different visualisation techniques are required in order to navigate and compare hierarchies with such different properties.

4. Conclusions and further work

The Nomencurator project is work in progress. There are several issues that remain to be addressed in the visualisation component of the project, the TaxoNote Comparator. These include:

- Extension of the visualisation and comparison to more than two trees. Indeed, a taxonomic revision of the Cryptomonads (in which some of the authors have particular expertise) has no less than eleven hierarchies that have been proposed in the last 100 years of work on the group [Novarino and Lucas 1995]. It is worth noting here that the current implementation accepts and integrates more than two hierarchies.

- Navigation over hierarchies. The Nomencurator data structure supports a data object called the Annotation which interconnects taxon concepts in multiple hierarchies. This documents statements made by the authors of a publication, such as "these names are synonyms for the following reason". The Assignment table is the place to show such data as a pop-up when one taxon in a hierarchy selected, and the mouse cursor is moved to another name in the table.

- Although the Nomencurator data structure was developed for biological taxonomy, the Comparator can be applied to other, more general areas of computing such as mapping between XML schemas, or ontologies.

References

BEDERSON, B. B., SHNEIDERMAN, B. AND WATTENBERG, M. 2002. Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies, *ACM Transactions on Graphics 21*, 4, 833 - 854.

BERENDSOHN, W. G. 1995. The concept of potential taxa in databases, *Taxon 44*, 2, 207-212.

BERENDSOHN, W. G. 1997. A taxonomic information model for botanical databases: the IOPI model, *Taxon 46*, 2, 283-309.

GANTER, B. AND WILLE, R., 1999. *Formal Concept Analysis: Mathematical Foundations,* Springer-Verlag.

GRAHAM, M. AND KENNEDY, J. 2001. Combining linking & focusing techniques for a multiple hierarchy visualisation. In *Fifth International Conference on Information Visualisation*, IEEE Computer Society Press. 425-432.

JUNG, S., PERKINS, S., ZHONG, Y., PRAMANIK, S. AND BEAMAN, J. 1995. A new data model for biological classification, *Computer Applications in the Biosciences 11*, 3, 237-246.

MUNZNER, T., GUIMBRETIÈRE, F., TASIRAN, S., ZHANG, L. AND ZHOU, Y. 2003. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. In *ACM SIGGRAPH*, ACM Press.

NOVARINO, G. AND LUCAS, I. A. N. 1995. A Zoological Classification-System of Cryptomonads, *Acta Protozoologica 34*, 3, 173-180.

PULLAN, M. R., WATSON, M. F., KENNEDY, J. B., RAGUENAUD, C. AND HYAM, R. 2000. The Prometheus Taxonomic Model: a practical approach to representing multiple classifications, *Taxon 49*, 1, 55-75.

RIDE, W. D. L., COGGER, H. G., DUPUIS, C., KRAUS, O., MINELLI, A., THOMPSON, F. C. AND TUBBS, P. K., 1999. *International code of zoological nomenclature : adopted by the International Union of Biological Sciences,* International Trust for Zoological Nomenclature.

YAO, Y. Y., WONG, S. K. M. AND T.Y.LIN. 1997. In *Rough Sets and Data Minig Analysis of Imprecise Data* (Eds, T.Y.Lin and N.Cercone) Kluwer Academic Publishers, pp. 47-75.

YTOW, N., MORSE, D. R. AND ROBERTS, D. M. 2001. Nomencurator: a nomenclatural history model to handle multiple taxonomic views, *Biological Journal of the Linnean Society 73*, 1, 81-98.