CHAPTER 6.1

Summary Databases and Model Repositories

Michael A. Arbib[1] and Amanda Bischoff-Grethe[2]

[1] University of Southern California Brain Project, University of Southern California, Los Angeles, California

[2] Center for Cognitive Neuroscience, Dartmouth College, Hanover, New Hampshire

Abstract

A *Summary Database* is like a review article but is structured as entries in a database rather than as one narrative, storing such high-level data as assertions, summaries, hypotheses, tables, and figures that encapsulate the "state of knowledge" in a particular domain. A *Model Repository* is a database that not only provides access to computational models but also links each model to the Empirical and Summary Databases to provide evidence for hypotheses in the model or to data to test predictions from simulation runs made with the model. This chapter provides a conceptual framework for databases of these types, pointing to specific instances of such databases—one implicit in the structure of our Brain Models on the Web (BMW) Model Repository as well as the NeuroScholar and NeuroHomology databases—described in later chapters and setting an agenda for future research and development.

6.1.1 The Database Typology and the NeuroInformatics Workbench

In Chapter 1.1, we introduced a fourfold typology of types of databases of relevance to Neuroinformatics: Article Repositories, Repositories of Empirical Data, Summary Databases, and Model Repositories. In this section, we briefly review what the earlier chapters of this volume have shown concerning our approach to providing tools for the construction and use of such databases. In particular, we will see the relationship of the key components of our NeuroInformatics Workbench—NeuroCore, NeuARt, NSLJ, and Annotator—to this typology. The rest of the chapter will provide a conceptual framework for the specific Summary Databases (see Chapters 6.3 and 6.4) and Model Repository (Chapter 6.2) that we have created to date, emphasizing not only our current status but also the long-term goals that have informed our work.

Article Repositories

These are in part the domain of commercial publishers and scientific societies, but can also take the form of resources constructed for access by some community over the Web or as a repository for technical reports for some organization. As such, we can leave the vigorous development of the tools for creating and using such repositories to others, but the USCBP has developed a simple prototype Model of Online Publishing (Chapter 5.3) to show how to increase the utility of online journals, etc. by offering new ways to link them to Repositories of Empirical Data and personal databases.

Repositories of Empirical Data

Our approach to Repositories of Empirical Data has emphasized the notion of a protocol for each experiment for which data are stored. The *protocol* provides information on the hypotheses being tested, the experimental methods used, etc. We have developed NeuroCore (Part 3) as our basic design for such databases, providing a general data schema which neuroscientists can extend readily to provide a tailored data structure that is still easy to understand by investigators from other laboratories. NeuroCore is enriched by the NeuroCore Time-Series Datablade (Appendix A2) to support storage of neuroscience time-series data. We further extended the utility of such repositories by developing NeuARt (Chapter 4.3) and NeuroSlicer (Chapter 4.4) to support the registration of data against standard brain atlases, with implications for database structuring and data retrieval.

Summary Databases

A *Summary Database* is like a review article but is structured as entries in a database rather than as one narrative, storing such high-level data as assertions, summaries, hypotheses, tables, and figures that encapsulate the "state of knowledge" in a particular domain. Our three USCBP Summary Databases, one implicit in the structure of our Brain Models on the Web (BMW) Model Repository, as well as the NeuroScholar[1] and NeuroHomology databases, will be described in subsequent chapters. We have already presented Annotator, our annotation technology (Chapter 5.4) for building a database of annotations on documents and databases scattered throughout the Web. The key observation is that a Summary Database can be seen as a form of annotation database with each summary serving as an annotation on all the clumps (selected items) that it summarizes. However, our work on BMW, NeuroScholar, the NeuroHomology database, and Annotator to date have proceeded in parallel and are only just now reaching the stage where we can begin to implement the conceptual integration that has been a foundation for our work.

Model Repositories

We envision a *Model Repository* as a database that not only provides access to computational models but also links each model to Empirical and Summary Databases to provide evidence for hypotheses in the model or to data to test predictions from simulation runs made with the model. Chapter 6.3 presents the current status of our Model Repository BMW (Brain Models on the Web). At present, BMW primarily includes models written in various versions of our Neural Simulation Language NSL (Chapter 2.2); future versions will reach all the way from models of brain imaging (Chapter 2.4) through NSL models to detailed compartmental models (NEURON and GENESIS) down to the finest levels of macromolecular structure (as in some of the EONS modeling of Chapter 2.3). We have already noted (Chapter 2.1) the utility of providing the implementation of a model with simulation interfaces which mimic experimental protocols so that operations on the model capture the manipulations the experiment might have made on corresponding portions of the nervous system.

6.1.2 An Overall Perspective

Our aim is to create an environment in which modelers and experimentalists can work together to increase our understanding of nervous systems. Fig. 1 is based upon the following points:

1. Laboratory data mean little unless they are stored in the database with information about the protocol used to obtain them.
2. Much of our access to such data will be indirect, via empirical generalizations and summary tables and figures.
3. Such summaries should be linked to laboratory data from which they are derived and/or journal articles in which they have been published.
4. Tools are needed to summarize laboratory data for publication.
5. Tools are needed to contrast and compare experimental results, generalizations, and the predictions of models of neural activity.

When working in neuroscience, a multitude of questions may occur regarding the data: "Is there a projection in the rat from the substantia nigra pars compacta to the subthalamic nucleus? How strong is this projection? Has this projection been modeled?" These and other questions can lead to time-consuming searches through published articles retrieved via such services as PubMed or Current Contents. First, there is the problem of knowing the appropriate keywords to retrieve the data relevant to the search; second, each article must be examined to determine if it does, indeed, cover the topic in question; and, finally, the article must be read more closely to find the sentence, paragraph, or figure that adequately answers the question. Although the answer is known for now, what about several weeks or months later when the material is needed again? Much of the same process is repeated, only this time the search is through a

[1] The NeuroScholar project was initiated as part of the USC Brain Project and is now being further developed by Gully Burns as a separate but federated research project.

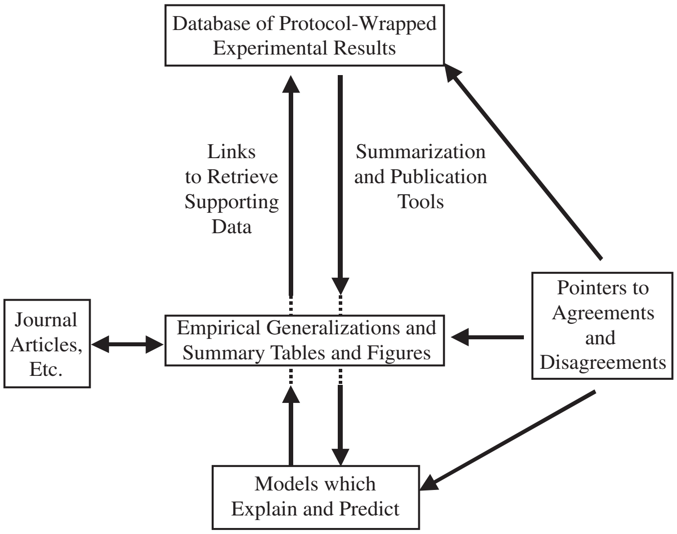

Database of Protocol-Wrapped Experimental Results
Links to Retrieve Supporting Data
Summarization and Publication Tools
Journal Articles, Etc.
Empirical Generalizations and Summary Tables and Figures
Pointers to Agreements and Disagreements
Models which Explain and Predict

Figure 1 A functional perspective on the integration of diverse databases in the USC Brain Project framework. The figure stresses the task of relating experimental data to empirical generalizations and the assumptions and predictions of computational models of neural function.

pile of papers on one's desk instead of through a search engine, culminating in the search for the paragraph that may or may not have been marked to indicate an important point. These summary statements that support our data can occur again and again and can necessitate a document search again and again, leading to time lost and often frustration when the answer does not immediately appear. Using a database to store this material, then, may be the answer to faster research and better organization of the data most important to one's work. The reader will note here a resonance with the general issues discussed in our introduction to annotation databases (Chapter 5.4).

A Summary Database (SDB), then, is a repository for summary data, documents, and annotations. A summary is a brief statement or description that summarizes a more complex body of text (or graphics or tables or graphs, etc.). For example, several different documents may give a precise description of a methodology used to determine connectivity between two regions and the results obtained by applying it. The summary data, then, would state that the connection does (or does not) exist, linked to the precise description from the source data, or "clumps" of the original document. In this fashion, one may search for generalities (i.e., does region A project to region B) and them examine more closely the source to help form a personal opinion on the summarized data.

Data stored within the SDB may relate to various aspects of neuroscience. The most obvious type is neuroanatomical data describing a cell group, receptor, channel, neurochemical, and so on. Neuroanatomical data, though, may be biological or simulated. A model stored within BMW (see Chapter 6.2 for details) may describe relations among several cell groups. For example, a model of the basal ganglia may state that a population of neurons, called put_inh, project to modeled external globus pallidus (GPe) neurons. When one does a search of summaries, then, one may choose to search for biological data, simulated data, or both. Storing these kinds of data together—whether in a single database or the databases of a well-designed federation—makes it easy to search for support (or inconsistencies). For example, a model of the basal ganglia in monkey may include a projection from the substantia nigra pars compacta (SNc) to the subthalamic nucleus (STN). Biologically, this projection is not known to exist in monkey; however, it does exist in the rat.

A summary statement may be a simple statement, such as region A projects to region B, or may contain more detailed information, such as the description of a cell type's dendritic tree or the pattern of connectivity between two regions. These data may serve further as support for a homology (Chapter 6.4). For example, a homology based upon relative position will provide the details regarding the relative position as a summary; another homology based upon connectivity will provide details regarding the afferents and/or efferents to the region and the pattern of connectivity. Alternatively, the user may view the originating text(s) supporting a summary statement.

In the SDBs we have developed in USCBP to date, we have stored the supporting data, called "*clumps*," for summaries directly within our databases. These clumps have been primarily portions of text from various documents, usually published research articles, but one can include figures, extracts from tables, etc. In Chapter 5.4 on annotations (a summary may be viewed as an annotation on the clumps that it summarizes), we have presented our software for creating and using Extended URLs to locate clumps from documents on the Web, not only retrieving the document but also highlighting the clump of interest. It is a challenge to extend the Extended URL methodology to characterize other types of clumps, such as subtables of a database or extracts from graphics, videos, or simulation runs. When using this technology, only the Extended URL, rather than the clump itself, need be stored in the Summary Database. This is important either when the clump is larger than a "fair use" portion of copyrighted material or when the clump is a very large object (e.g., recordings from multiple electrodes, video clips, etc.). The latter consideration reminds us of the mass storage issues discussed in Chapter 5.5.

Another general issue for SDBs is that neuroscience data do not have the unequivocal nature of entries such as, say, "Mr. X has seat Y on flight W on date D" in an airline reservation database. A seemingly clear statement such as "Region A connects to Region B" may depend on the protocol used, variability between animals within and across species, the skill of the experimenter, and the specific interpretation of how best to characterize the two regions. Thus, a summary (like a review article) is itself the result of both the choice of clumps to summarize and of the collator's choices in synthesizing what may be apparently discordant material. Both Chapters 6.3 (NeuroScholar) and 6.4 (NeuroHomology) pay attention to these issues. Each summary should contain a field naming the collator; each summary should come not only with copies of, or pointers to, the clumps that are summarized, but also—where possible—with a rationale for the summary and a confidence level (between 0 and 1) for the assertion that the summary provides. Moving beyond the current state of our art, we envisage the following.

Summaries may be either unrefereed or refereed. In any case, summaries will be annotated with new data of relevance, critiques, annotations, etc. New search tools (such as those planned for our annotation technology, Chapter 5.4) will make it easy to use an existing summary to anchor a search for related data. The pace of research means that summaries will need constant updating—even if their general conclusions remain correct, new data may markedly change the confidence levels for the assertions they contain. When a critical mass of such changes has accumulated, or when a synthesis of related

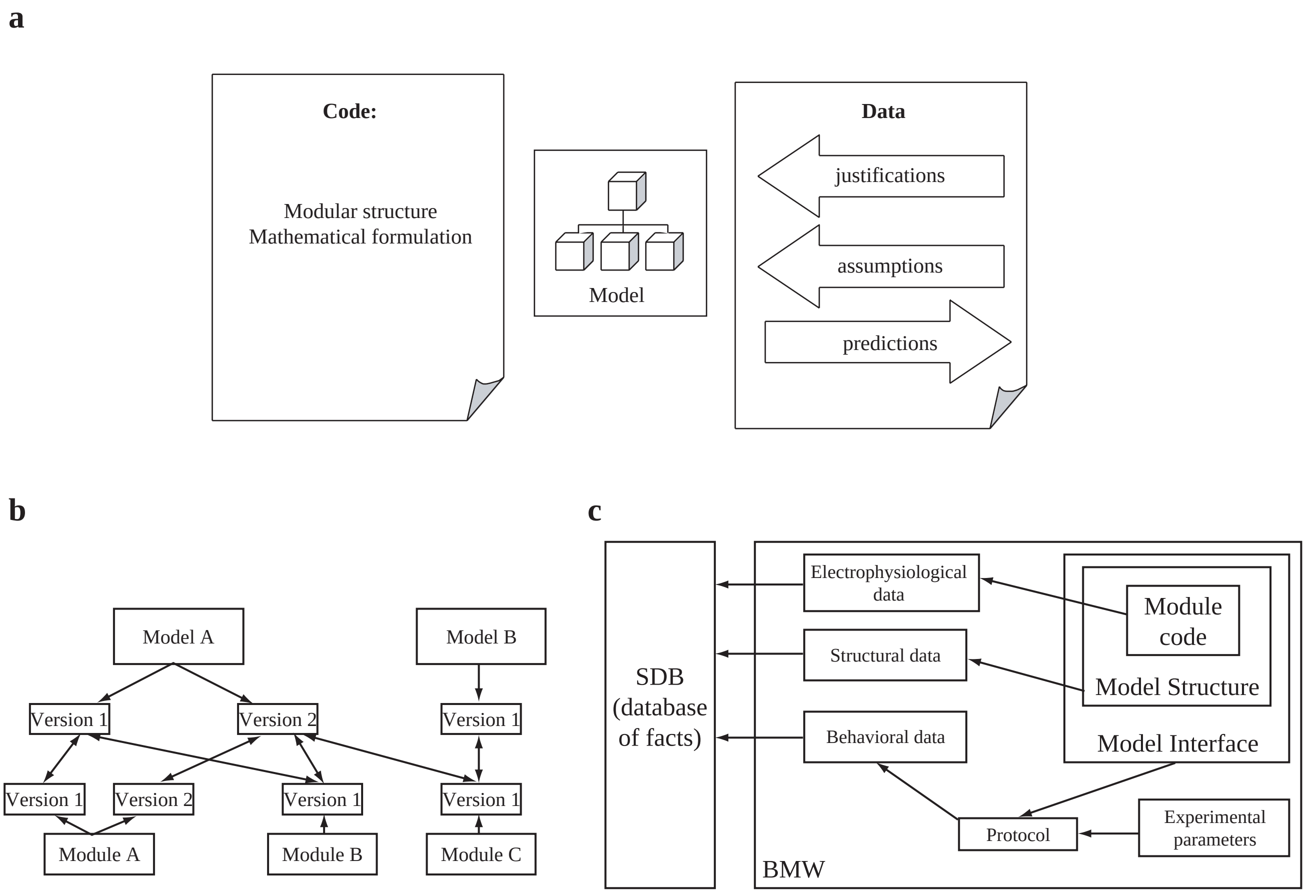


Figure 2 **(a)** Overview of two views of a model. **(b)** More detailed schematic of the modular structure of the model, emphasizing that a model, and modules, may have different versions; shows how links to a Summary Database may be used to test and/or justify assumptions and predictions of the model. **(c)** Emphasizes the other dimension of a Model Repository, documenting the relation of the model to empirical data and (although this is not made explicit in the diagram) other models (see also Fig. 1).

summaries is undertaken, then a new summary can be posted which can be seen as a new version of the one it replaces. As in all cases of versioning, there will be the editorial decision regarding refereed databases as to when a new version becomes the “official” version, in which case the old versions should still be available for (rare) consultation, but perhaps in an archive based on a slower but less expensive mass storage medium (*cf.* the issues discussed in Chapter 5.5).

6.1.3 General Considerations on Model Repositories

Fig. 2 presents our general perspective on Model Repositories. On the one hand (part b), the modeler will want access to the model in all its detail, ranging from viewing the modular structure at various levels of detail all the way down to the code. The repository should also allow access to different versions of each model and model. As emphasized in Chapter 2.2, we have developed the Schematic Capture System (SCS) to ease the creation and viewing of the modular structure of modules written in NSL. It should be a straightforward extension to enable the SCS to display any and all modular models. The greater challenge (one to be shared with other groups beyond USCBP) will be to ensure the *interoperability* of models developed using different simulators, “wrapping” modules so they can work together, thus allowing the creation of new versions of old models, and new models as well. A current project is to develop SCS so that it not only allows both the modeler and experimentalist to view the components of a model and their connections, with links to the supporting NSL code, but will also allow one to follow links to summary data that test or justify the model assumptions expressed in the SCS graphics.

We also show in Fig. 2 the SDB as a database separate from the Model Repository (in this case, BMW); however, in the current prototype of BMW, the SDB is a subset of the BMW database, rather than a separate database. In the remainder of this chapter, we briefly review the structure of the SDB contained within the current implementation of BMW. We then introduce key features of our other two current USCBP Summary Databases, NeuroScholar and NeuroHomology. Further details of these three SDBs are provided in Chapters 6.2, 6.3, and 6.4, respectively.

6.1.4 Brain Models on the Web

Brain Models on the Web (BMW) is a Model Repository which allows users to run experiments, build new models, and keep track of versions within a database. SDB and BMW are constructed to work together and take advantage of material that each database stores. Fig. 3 describes the relationship between the two databases. A researcher creating a model within BMW will rely upon SDB for data to support his model assumptions and test predictions and will tie these summaries to the appropriate areas of the model. When the model is complete, he may choose to store several simulation runs of the different neural populations represented within the model. These simulations may also be documented within SDB. Similarly, SDB may hold information on the model’s protocols and interface. Another researcher working in SDB, then, may perform a search on modeling data and be returned information regarding the model in BMW and its supporting data. He may then decide to run the model himself, perhaps trying slightly different protocols, to validate that the neural population within the model does indeed match expectations based upon available literature. This researcher may choose to take this one step further and review side-by-side comparisons of the simulated model and real biological data stored within NeuroCore.

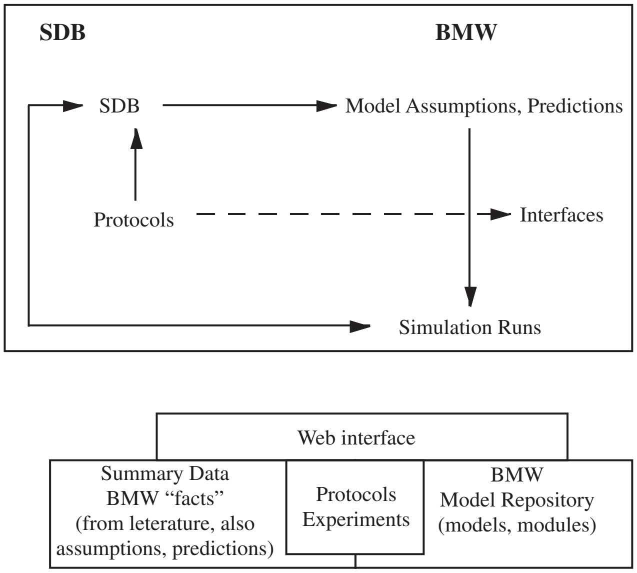
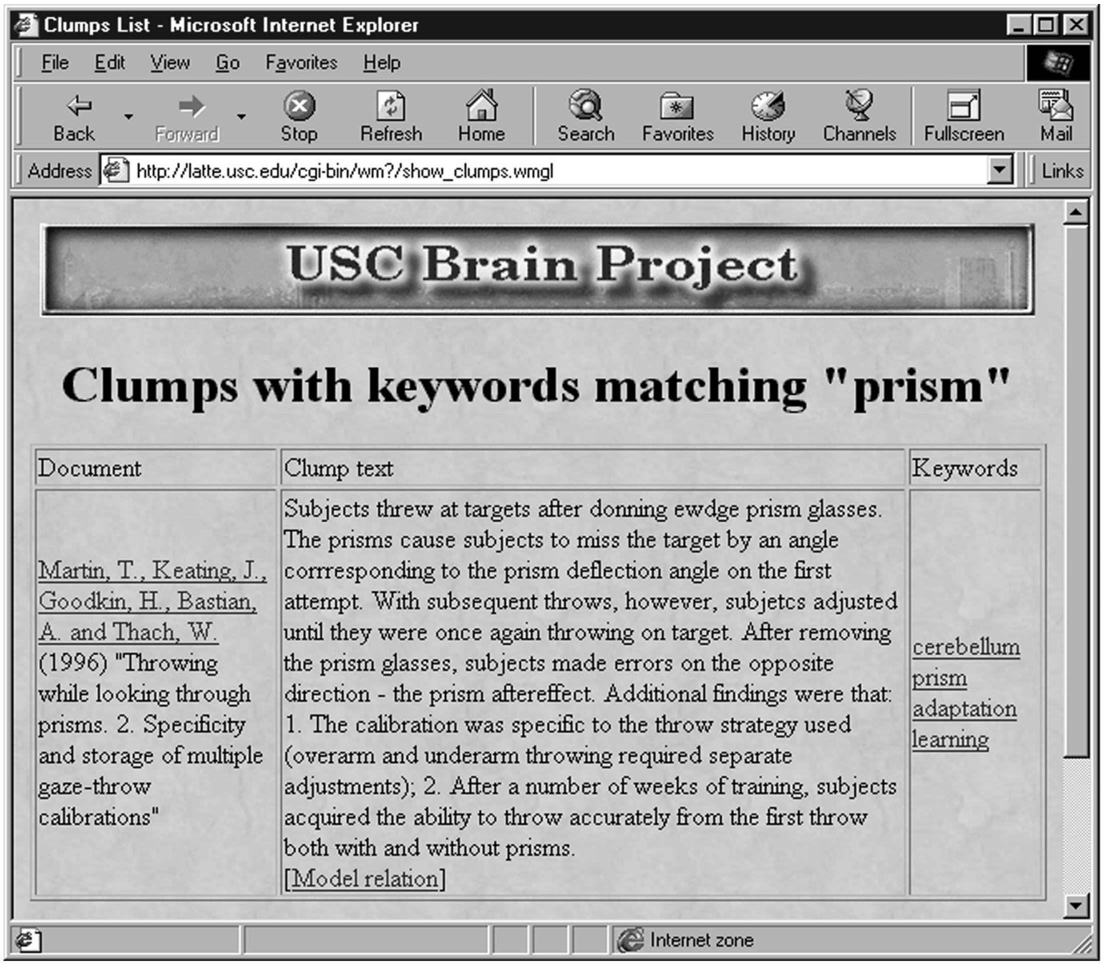


Figure 3 A diagram showing the relationship between SDB and BMW. Summary statements within SDB can be used to form model assumptions and predictions. The completed model and its simulation runs can be documented via SDB to allow researchers to access both biological and simulated data on the same topic. In the current implementation of BMW, the Summary Database is embedded within BMW as a set of facts that can be related to modules, rather than being a separate database or federation of Summary Databases.

As already noted, the current BMW contains an embedded Summary Database (Fig. 3); details are provided in Chapter 6.2. Here, we briefly note the way in which clumps are handled in the SDB of the current BMW. The entry for a clump contains the text of the clump (as distinct from Extended URLs pointing to clumps; see Chapter 5.4), together with the reference to the document that contains the clump and a set of keywords. Fig. 4 shows a clump retrieved by a search on the keyword “prism.” The current version does not contain explicit summaries; rather, it stores relations between



Address: http://latte.usc.edu/cgi-bin/wm?/show_clumps.wmgl

USC Brain Project

Clumps with keywords matching "prism"

Document	Clump text	Keywords
Martin, T., Keating, J., Goodkin, H., Bastian, A. and Thach, W. (1996) "Throwing while looking through prisms. 2. Specificity and storage of multiple gaze-throw calibrations"	Subjects threw at targets after donning ewdge prism glasses. The prisms cause subjects to miss the target by an angle corrresponding to the prism deflection angle on the first attempt. With subsequent throws, however, subjetcs adjusted until they were once again throwing on target. After removing the prism glasses, subjects made errors on the opposite direction - the prism aftereffect. Additional findings were that: 1. The calibration was specific to the throw strategy used (overarm and underarm throwing required separate adjustments); 2. After a number of weeks of training, subjects acquired the ability to throw accurately from the first throw both with and without prisms. [Model relation]	cerebellum prism adaptation learning

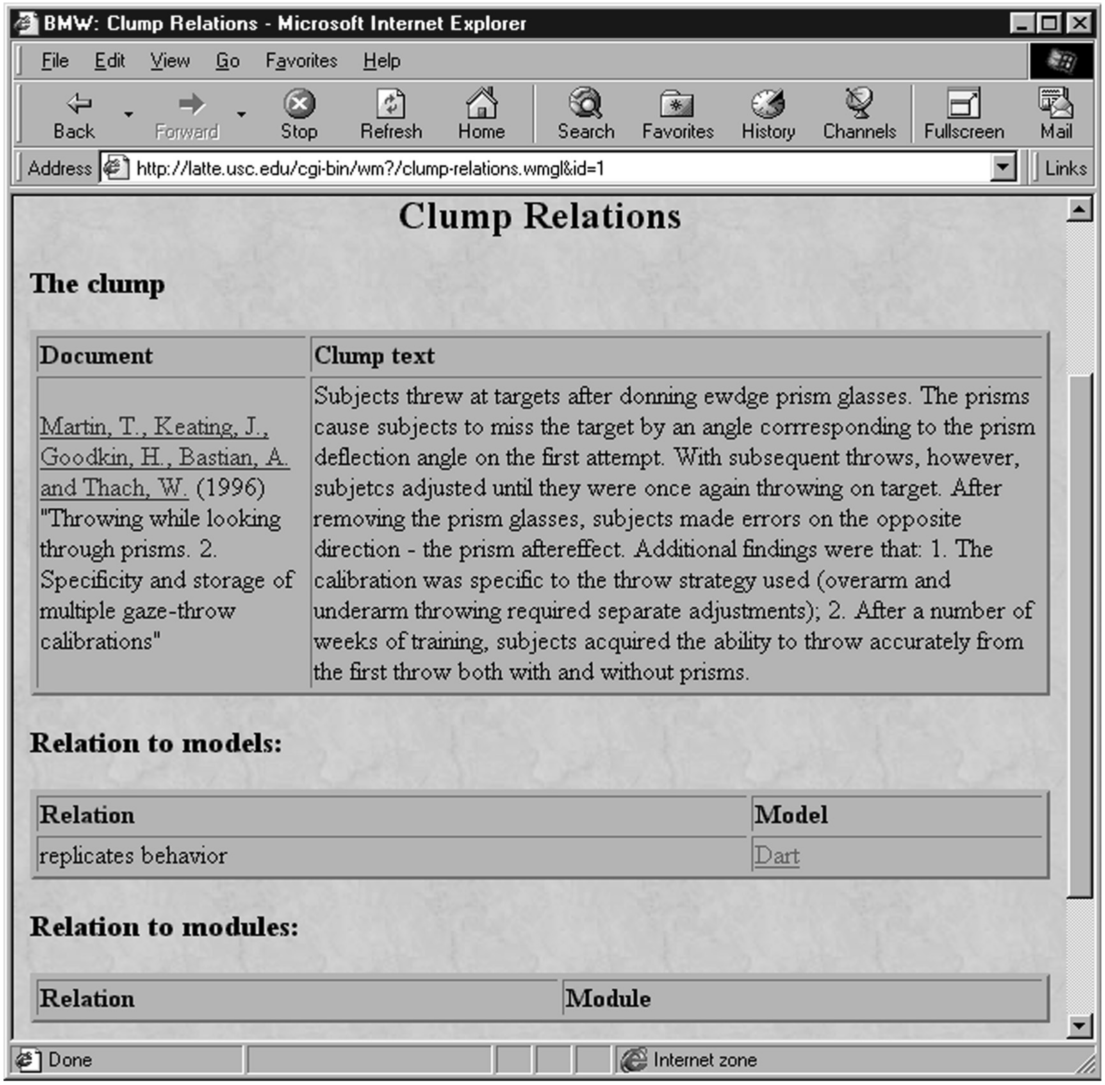
Figure 4 The current version of BMW contains an SDB which contains clumps. The entry for the clump contains the text of the clump (as distinct from Extended URLs pointing to clumps; see Chapter 5.4) together with the reference to the document that contains the clump and a set of keywords. Here we see a clump retrieved by a search on the keyword "prism."

clumps expressing empirical data and aspects of the model—whether model assumptions or simulation results—stored in BMW. Thus, a search can reveal a summary implicit in all these links. For example, given a clump, we can search for modules and models to which it is related. Fig. 5 gives an example relating a clump to the "Dart" model via the relation "Replicates behavior." Fig. 6 shows some of the clumps retrieved by querying which clumps are related to a given model. Note that the response is related to the hierarchical structure of the model—we see clumps related to the top-level description of the model as well as model behavior; we also see clumps related to the constituent modules of the model. At present, the set of such relations is relatively limited, and the search for such relations is still conducted by hand.

6.1.5 NeuroScholar

NeuroScholar is a Summary Database, that is, a knowledge-base management system for neuroscientific information taken from the literature. The key issue that drives the design of this SDB is to allow users to compile, interpret, and annotate low-level, textual descriptions of neuroanatomical data from the literature (Chapter 6.3). NeuroScholar is founded on the following conceptual entities (Fig. 7):

1. *Atoms:* These are raw text and figures that can be found in the literature (we call these clumps elsewhere).
2. *Primitives:* These are classifications of atoms so that the information they are concerned with can be structured in a useful way. Here, we differentiate between "knowledge types" representing qualitative differences between data based on data from a paper's abstract or results section as author interpretations, and descriptions. The properties of a given primitive returns subprimitives, allowing the construction of complex data entities that closely correspond to neuroscientific concepts found in the literature.
3. *Relations:* These are used to store rules that compare and link different primitives. This is especially useful


BMW: Clump Relations - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Channels
Fullscreen
Mail
Address http://latte.usc.edu/cgi-bin/wm?/clump-relations.wmgl&id=1
Links
Done
Internet zone

Clump Relations

The clump

Document	Clump text
Martin, T., Keating, J., Goodkin, H., Bastian, A. and Thach, W. (1996) "Throwing while looking through prisms. 2. Specificity and storage of multiple gaze-throw calibrations"	Subjects threw at targets after donning ewdge prism glasses. The prisms cause subjects to miss the target by an angle corrresponding to the prism deflection angle on the first attempt. With subsequent throws, however, subjetcs adjusted until they were once again throwing on target. After removing the prism glasses, subjects made errors on the opposite direction - the prism aftereffect. Additional findings were that: 1. The calibration was specific to the throw strategy used (overarm and underarm throwing required separate adjustments); 2. After a number of weeks of training, subjects acquired the ability to throw accurately from the first throw both with and without prisms.

Relation to models:

Relation	Model
replicates behavior	Dart

Relation to modules:

Relation	Module

Figure 5 Given a clump, we can search for modules and models to which it is related. Here we see an example relating the clump to the "Dart" model via the relation "Replicates behavior." At present, the set of such relations is relatively limited, and the search for such relations is still conducted by hand.

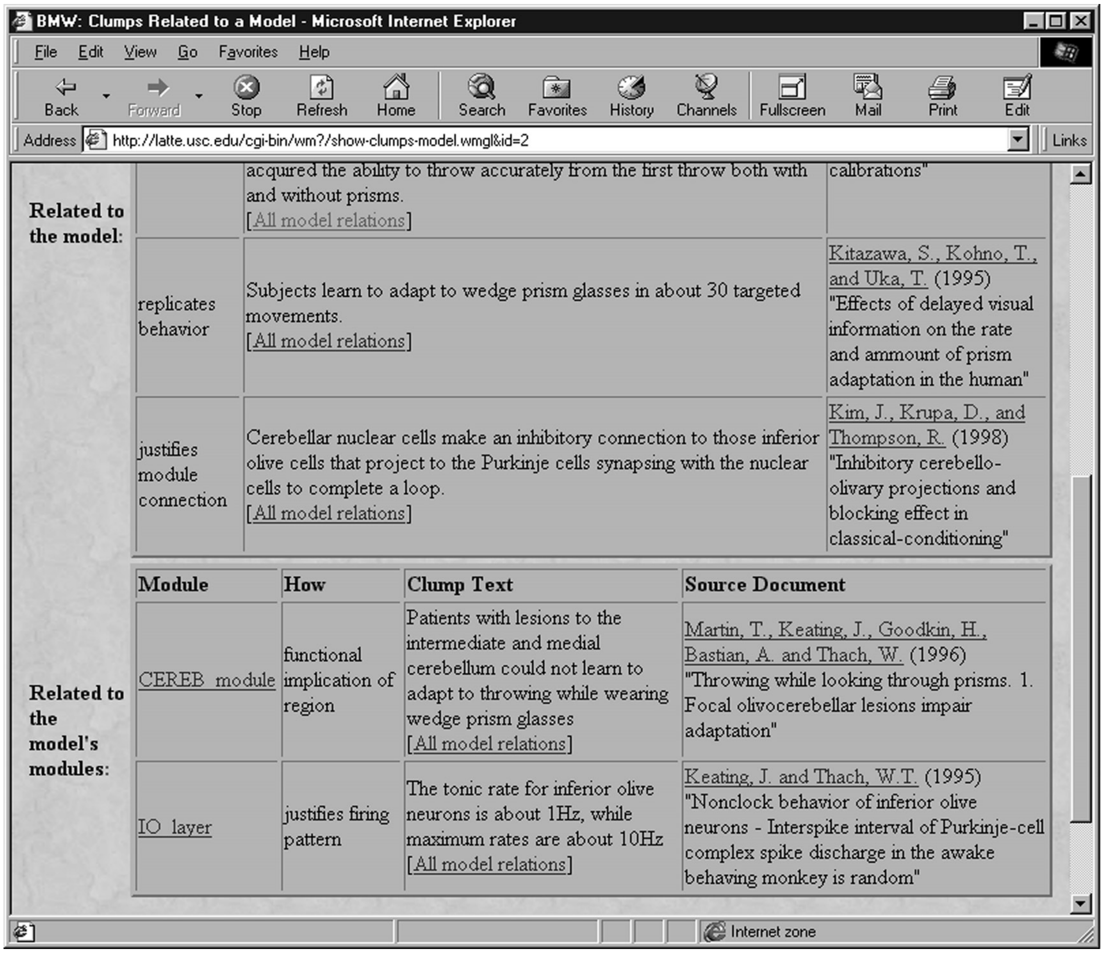
when linking primitives that are derived from different papers.

4. *Annotations:* NeuroScholar allows users to make two sorts of annotations—comments, which are just remarks that users wish to attach to atoms, publications, primitives, or relations, and justifications, which are required to substantiate any relation or primitives that are classified as user interpretations.

Fig. 8 shows a number of examples of NeuroScholar Primitives: *brainVolumes* are simply volumes of brain tissue possessing geometrical properties; *blackBoxConnections* represent neural connections and have two subsidiary object: the *somata brainVolume* denotes the region where the connection originates and the *terminalField brainVolume* denotes the region where the connection terminates.

In Fig. 8b, we add more advanced NeuroScholar primitives: *neuronPopulations* are populations of cells, grouped on the basis of a stated criterion; *dendrites, somata, axons,* and *terminalField* objects represent shared properties of the constituent neurons of *neuronPopulation* objects. The properties of these subprimitives could be defined to accommodate a range of different properties from different papers (such as the presence of labeled mRNA, the morphology of the somata, or firing characteristics of neurons).

We intend to combine NeuARt, our atlas-based user interface for a NeuroCore database that handles neuroanatomical data, and NeuroScholar so that specified brainVolumes can be linked to volumes embedded in the relevant brain atlas (*cf.* Chapters 4.3 and 4.4 for the relevant context for this ongoing work). This means that the framework for published data can be used to structure and interpret experimental data while also linking the geometrical properties of the of published data to an atlas. This exemplifies the spatial query manager concept


BMW: Clumps Related to a Model - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Channels
Fullscreen
Mail
Print
Edit
Address http://latte.usc.edu/cgi-bin/wm?/show-clumps-model.wmgl&id=2
Links
Internet zone

Related to the model:

	acquired the ability to throw accurately from the first throw both with and without prisms. [All model relations]	calibrations"
replicates behavior	Subjects learn to adapt to wedge prism glasses in about 30 targeted movements. [All model relations]	Kitazawa, S., Kohno, T., and Uka, T. (1995) "Effects of delayed visual information on the rate and ammount of prism adaptation in the human"
justifies module connection	Cerebellar nuclear cells make an inhibitory connection to those inferior olive cells that project to the Purkinje cells synapsing with the nuclear cells to complete a loop. [All model relations]	Kim, J., Krupa, D., and Thompson, R. (1998) "Inhibitory cerebello-olivary projections and blocking effect in classical-conditioning"

Related to the model's modules:

Module	How	Clump Text	Source Document
CEREB module	functional implication of region	Patients with lesions to the intermediate and medial cerebellum could not learn to adapt to throwing while wearing wedge prism glasses [All model relations]	Martin, T., Keating, J., Goodkin, H., Bastian, A. and Thach, W. (1996) "Throwing while looking through prisms. 1. Focal olivocerebellar lesions impair adaptation"
IO layer	justifies firing pattern	The tonic rate for inferior olive neurons is about 1Hz, while maximum rates are about 10Hz [All model relations]	Keating, J. and Thach, W.T. (1995) "Nonclock behavior of inferior olive neurons - Interspike interval of Purkinje-cell complex spike discharge in the awake behaving monkey is random"

Figure 6 Here we see some of the clumps retrieved by querying which clumps are related to a given model. Note that the response is related to the hierarchical structure of the model; we see clumps related to the top-level description of the model as well as model behavior; we also see clumps related to the constituent modules of the model.

of NeuARt—that relevant neuroscience data can be based on geometrical queries to an atlas, as well as on text-based and other forms of query.

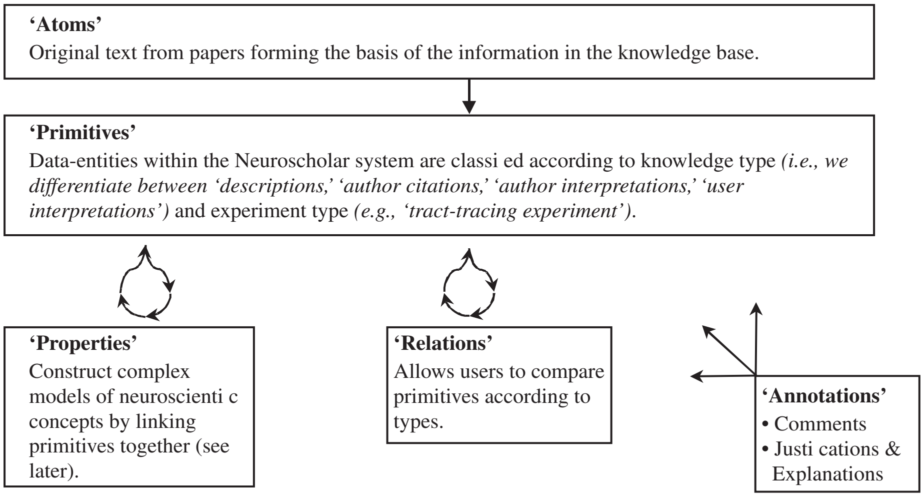
Chapter 6.3 provides much more information on NeuroScholar, including a critical discussion of the evaluation data used to form summaries and an indication of the data analysis tools that have been used to mine interesting conclusions from connectivity data. NeuroScholar's weighted scheme for "believability" in connectivity provides a useful model for evaluating data summaries for other areas of neuroscience.

6.1.6 NeuroHomology

The NeuroHomology database (Chapter 6.4) is a knowledge-based Summary Database that is designed to aid the search for homologies between brain structures for human/monkey, human/rat, and human/rodent species. To support this, the database contains searchable entries for brain structures and connections as interpreted from the literature, with quantification of the degrees of confidence of connections and staining techniques that are used. As in NeuroScholar (Chapter 6.3), we also propose a method of computing an overall confidence level for a number of entries related to a single connection, and based on it the user can evaluate the reliability of the searched connection as reflected in the literature.

We have identified eight criteria to define a homology between two brain structures: cell morphology, relative position, cytoarchitecture, chemoarchitecture, myeloarchitecture, afferent and efferent connections, and function. A "*homology inference engine*" lets one retrieve all data relevant to a possible homology between regions in different species and compute a degree of confidence that a homology does indeed exist.

The homologies part of the NeuroHomology database can handle data about brain structures from any species for which data are available. On the other hand, due to the fact that the homologies are seen as a tool for computational neuroscientists and we focus on models of brain structures from humans, monkeys, and rodents, the database contains human/rat, human/monkey, and



'Atoms'
Original text from papers forming the basis of the information in the knowledge base.

'Primitives'
Data-entities within the Neuroscholar system are classified according to knowledge type *(i.e., we differentiate between 'descriptions,' 'author citations,' 'author interpretations,' 'user interpretations')* and experiment type *(e.g., 'tract-tracing experiment')*.

'Properties'
Construct complex models of neuroscientific concepts by linking primitives together (see later).

'Relations'
Allows users to compare primitives according to types.

'Annotations'
- Comments
- Justifications & Explanations

Figure 7 NeuroScholar uses the term "atom" where we use the term "clump" but offers useful typologies not yet attempted in other work on Summary Databases in the USC Brain Project.

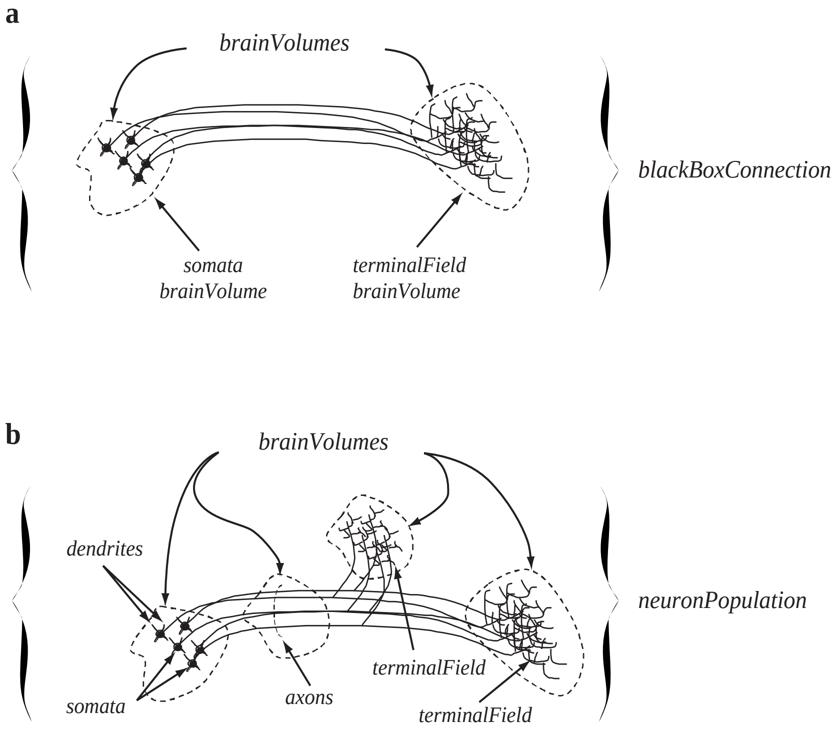
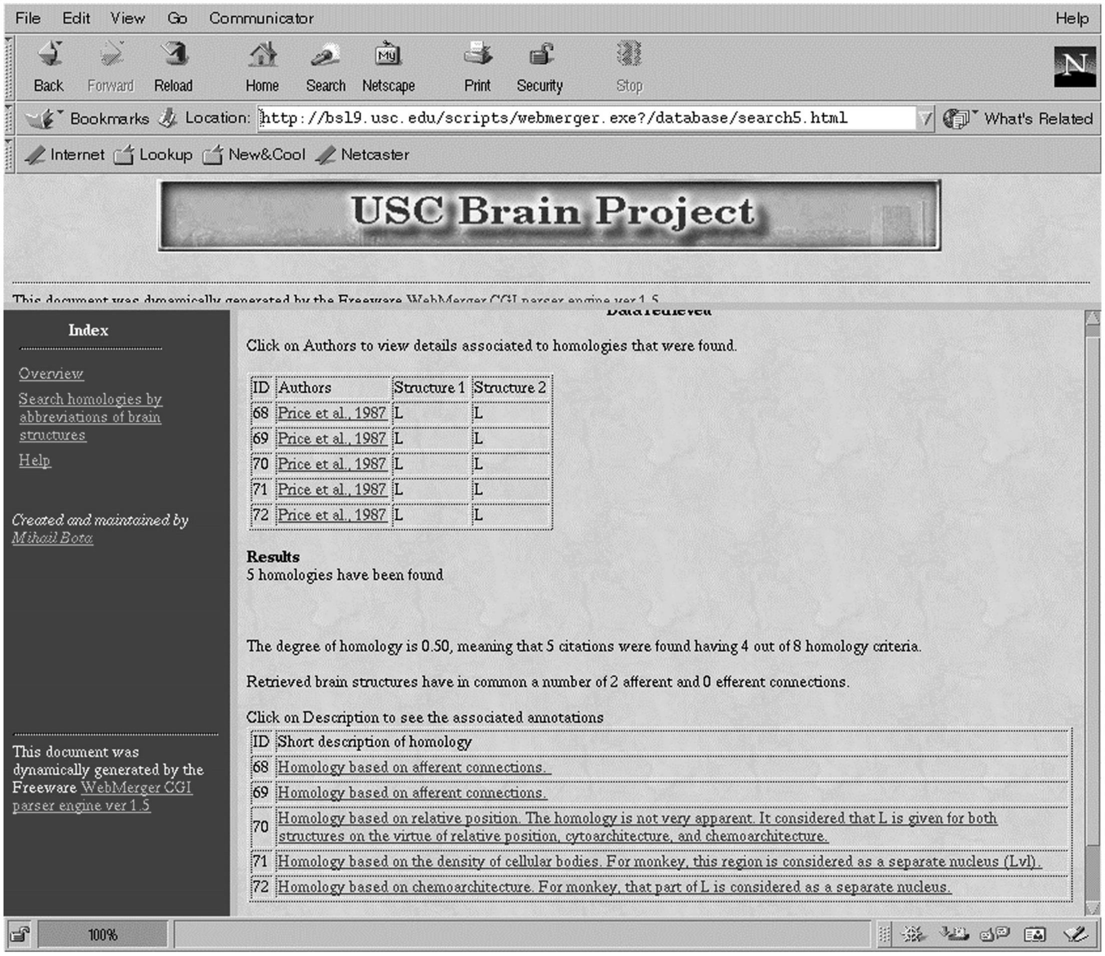


Figure 8 Examples of NeuroScholar primitives, with increasing detail in (a) as compared to (b).

monkey/rat homologies. The search of the NeuroHomology database can be made by abbreviations of brain structures. The user can enter the abbreviations for two brain structure from two different species (Fig. 9). The result of a search will retrieve all the homologies that are found in the database with regard to the searched pair of brain structures. The user can inspect all the details of any retrieved entry by clicking on the cited reference. The result will be eight different tables containing information about each homology criterion. Associated with each entry is a short description of the homology, which is inserted by the collator. The user can enter and inspect the additional annotations that are attached to each investigated entry. In this way, the user can evaluate the relative importance of common patterns of connectivity in evaluating the degree of homology. A second reason for showing the common afferent and efferent connections, as reflected from the inserted data in the database, is related to the functionality of brain areas. Two brain structures that share a common pattern of connectivity are likely to have the same position in the hierarchy of processing of information in the central



http://bs19.usc.edu/scripts/webmerger.exe?/database/search5.html

This document was dynamically generated by the Freeware WebMerger CGI parser engine ver 1.5

Index

Overview

Search homologies by abbreviations of brain structures

Help

Created and maintained by Mihail Bota

This document was dynamically generated by the Freeware WebMerger CGI parser engine ver 1.5

Data retrieved

Click on Authors to view details associated to homologies that were found.

ID	Authors	Structure 1	Structure 2
68	Price et al., 1987	L	L
69	Price et al., 1987	L	L
70	Price et al., 1987	L	L
71	Price et al., 1987	L	L
72	Price et al., 1987	L	L

Results
5 homologies have been found

The degree of homology is 0.50, meaning that 5 citations were found having 4 out of 8 homology criteria.

Retrieved brain structures have in common a number of 2 afferent and 0 efferent connections.

Click on Description to see the associated annotations

ID	Short description of homology
68	Homology based on afferent connections.
69	Homology based on afferent connections.
70	Homology based on relative position. The homology is not very apparent. It considered that L is given for both structures on the virtue of relative position, cytoarchitecture, and chemoarchitecture.
71	Homology based on the density of cellular bodies. For monkey, this region is considered as a separate nucleus (Lvl).
72	Homology based on chemoarchitecture. For monkey, that part of L is considered as a separate nucleus.

Figure 9 The query for homologies between monkey lateral nucleus of amygdala (L) and rodent L retrieved five different entries. The homology criteria that are fulfilled are afferent connections, relative position, cytoarchitecture, and chemoarchitecture (staining for AchE).

nervous system and have common functions. Moreover, the computational neuroscientists that use the NeuroHomology database can evaluate the degree of reliability to model neural systems from one species by using data from other species.

6.1.7 Future Plans

Although at present they are separate structures, we envision that the NeuroHomology database will be merged with the SDB currently embedded within BMW to form a generic SDB enriched with a "homology inference engine", as well as a number of features developed within the NeuroScholar project. We also plan to make full use of our annotation technology in the creation of SDBs and their linkage to BMW. On the more formal side, we are developing a set of key properties and relations to better relate neural structures to the model structures that represent them, enriching the anatomical examples coded in NeuroScholar with the rich set of constructs contained within the models already embedded in BMW. As noted, we are extending SCS from the representation of NSL models to the integration of models with empirical data and hope to engage the Neuroinformatics community in solving the interoperability problem for developing models, versioning, and linking them to empirical data in a modeling environment that contains a variety of simulators. Finally, much more needs to be done in developing criteria for comparing model simulations and predictions with empirical data, with special attention given to parameter identification techniques that will determine which settings of model parameters best match model performance to a given suite of test data.

Acknowledgments

This work was supported in part by the Human Brain Project (with funding from NIMH, NASA, and NIDA) under the P20 Program Project Grant HBP: 5–P20–52194 for work on "Neural Plasticity: Data and Computational Structures" (M. A. Arbib, Director). We thank Jacob Spoelstra, Gully A. P. C. Burns, and Mihai Bota for their work which enriched this chapter and which is reported in more detail in Chapters 6.2, 6.3, and 6.4, respectively.