Enhanced Free Text Access to Anatomically-Indexed Data

Gail Sinclair
Informatics Division
University of Edinburgh
80 South Bridge
Edinburgh, UK EH1 1HN
carols@dai.ed.ac.uk

Bonnie Webber
Informatics Division
University of Edinburgh
2 Buccleuch Place
Edinburgh, UK EH8 9LW
bonnie.webber@ed.ac.uk

Duncan Davidson
MRC Human Genetics Unit
Western General Hospital
Crewe Road
Edinburgh, UK EH4 2XU
duncan.davidson@hgu.mrc.ac.uk

Abstract

We describe our use of an existing resource, the *Mouse Anatomical Nomenclature*, to improve a symbolic interface to anatomically-indexed gene expression data. The goal is to reduce user effort in specifying anatomical structures of interest and increase precision and recall.

1 Introduction

Language Technology (LT) resources are time-consuming and expensive to develop, and applications rarely have the luxury of calling upon resources specially designed for the task at hand. For LT applications in developmental biology such as robust interfaces to anatomically-indexed gene expression data and text mining tools to assist in building such databases, resources already exist in the form of *anatomical nomenclatures* for several model organisms including mouse, zebrafish, drosophila and human. (Others may follow.) These nomenclatures have been developed by biologists for biologists, to record in a clear, intuitive and structured way the structures that can be distinguished at each stage of an embryo's development. The challenge for LT applications in developmental biology is to stretch them to serve other purposes as well.

In this paper, we describe how we have taken one of these anatomical nomenclatures (mouse) and extracted from it a new resource to facilitate free text access to anatomically-indexed data. The techniques we have used are applicable to anatomical nomenclatures for other model organisms as well.

The paper is organised as follows: In Section 2, we describe the Mouse Atlas, which is the particular context for the interface we are developing. Section 3 describes what we are doing to reduce the amount of effort a user has to expend in specifying anatomical structures of interest to them. In Section 4, we describe how what we did to reduce user effort also serves to provide a clearer display of the results of searching. Then in Sections 5 and 6, we describe what we are doing to increase the *precision* and *recall* of user queries.

2 The Mouse Atlas

The *Mouse Atlas*, developed by researchers at the Medical Research Council's Human Genetics Unit (MRC HGU) in Edinburgh, is a 3D atlas of mouse embryo development (http://genex.hgu.mrc.ac.uk). Anatomical structures within each of the 26 Theiler Stages of embryo development are labelled, and 3D reconstructions of each stage can be displayed in transverse, frontal, sagittal or arbitrary planes.

The *Mouse Atlas* is now being used to support indexing of gene expression data, allowing the results of gene expression experiments to be indexed with respect to where genes are expressed in the developing embryo. There are at least two ways of using anatomy to index gene expression data. In *spatial indexing*, data is associated directly with volume elements, *voxels* of the anatomical model. In *symbolic indexing*, gene expression data is associated with a label specifying a pre-defined region of the embryo.

A database of spatially indexed gene expression data (the EMAGE database) is being developed at the MRC HGU. A database of symbolically indexed gene expression data (the *Gene Expression Database* or GXD) is being developed by the Jackson Laboratory in Bar Harbor, Maine

Proceedings of the Workshop on Natural Language Processing in the Biomedical Domain, Philadelphia, July 2002, pp. 45-52.
Association for Computational Linguistics.

(http://www.informatics.jax.org). Indexing in the GXD uses the *Mouse Anatomical Nomenclature*, a set of 26 trees of anatomical terms (one tree per Theiler Stage) structured primarily by part-whole relations (and some set-member relations).

The root node of each Theiler Stage tree corresponds to the entire embryo at that stage, while other nodes correspond to organ systems, subsystems, spatially-localised parts of subsystems, or anatomical structures. Each node within a tree has a label (its *component term*), but *component terms are not meant to serve as unique designators*: the only thing guaranteed to denote an anatomical structure uniquely is the sequence of *component terms* that comprises a *path* from the root node. Thus paths (and only paths) can serve as keys for symbolic indexing of data. For example the component term CRANIAL labels both a child of GANGLION (i.e., ganglia located in the head) and a child of NERVE (i.e., nerves located in the head). The path to this latter child

EMBRYO.ORGANSYSTEM.NERVOUSSYSTEM.
CENTRALNERVOUSSYSTEM.NERVE.
CRANIAL.TRIGEMINALV

uniquely denotes the trigeminal, or fifth cranial, nerve[1] within the Theiler Stage denoted by its root and is used as a key for relevant data.

For *accessing* gene expression data, both spatial and symbolic means are again both possible. An elegant spatial interface is being completed at the MRC HGU, that pairs an active window containing a view of the embryo stage of interest, with a window containing the corresponding nomenclature tree[2]. Clicking at a point in the embryo view highlights the most specific corresponding node of the nomenclature being displayed (i.e., sub-trees of a node can be either hidden or exploded). Similarly, clicking on a term in the nomenclature highlights the corresponding structure within the embryo along the plane currently being displayed. A screen-shot from the interface is shown in Figure 2. On the left of the figure is an outline frontal drawing of the embryo, on which a sagittal section plane is marked in red. The centre panel shows a digital, sagittal section through the volumetric embryo model with the delineated left dorsal aorta coloured blue. The corresponding component term is highlighted on the Mouse Anatomical Nomenclature on the right. Users can access the gene expression data on the highlighted structure by another mouse click.

The current project aims at improving *symbolic* access to gene expression data by providing a robust free-text interface. In the current GXD interface[3], users can tree-walk through the *Mouse Anatomical Nomenclature* for a given stage, to find the anatomical structure whose associated gene expression data is of interest to them, or they can enter a term which is matched against *single* component names, with all possible substring matches returned for the user to choose among.

Problems exist with both forms of access. It is well known that navigating through a tree is tedious. A more subtle problem arises from anatomy being forced into a tree-structure that it doesn't have. This leads to structures being divided and the resulting sub-structures realised in different parts of the tree – for example, the endocardial tube is divided into one part that is a daughter of COMMON ATRIAL CHAMBER (i.e., its location), and another part which is a daughter of OUTFLOW TRACT. So appearing to find a structure of interest by a tree-walk, in addition to being tedious, doesn't by itself guarantee the user that s/he has found it all.

In contrast, access by sub-string matching on individual component terms has problems of both *recall* and *precision*. A user may enter a string that matches nothing (0% recall): (1) it may be neither a component term nor a substring within a component term – e.g., while HEART is a common synonym for the modifier "cardiac" (and a component term in its own right) and CARDIAC MUSCLE is a component term, the string "heart muscle" yields no match; or (2) it may span multiple component terms in the nomenclature – e.g., while GLAND is a component term, and PITUITARY is the component term of one of its children (i.e., a member of the set of glands), the string "pituitary gland" yields no match. Or the opposite may happen: 100% recall with low precision. For example, a search on "hindbrain" yields 22 matches, while "mesenchyme" yields as many as

[1] Full stop is used to separate component names along a path.
[2] http://genex.hgu.mrc.ac.uk/Resources/GXDQuery1
[3] http://www.informatics.jax.org/menus/expression_menu

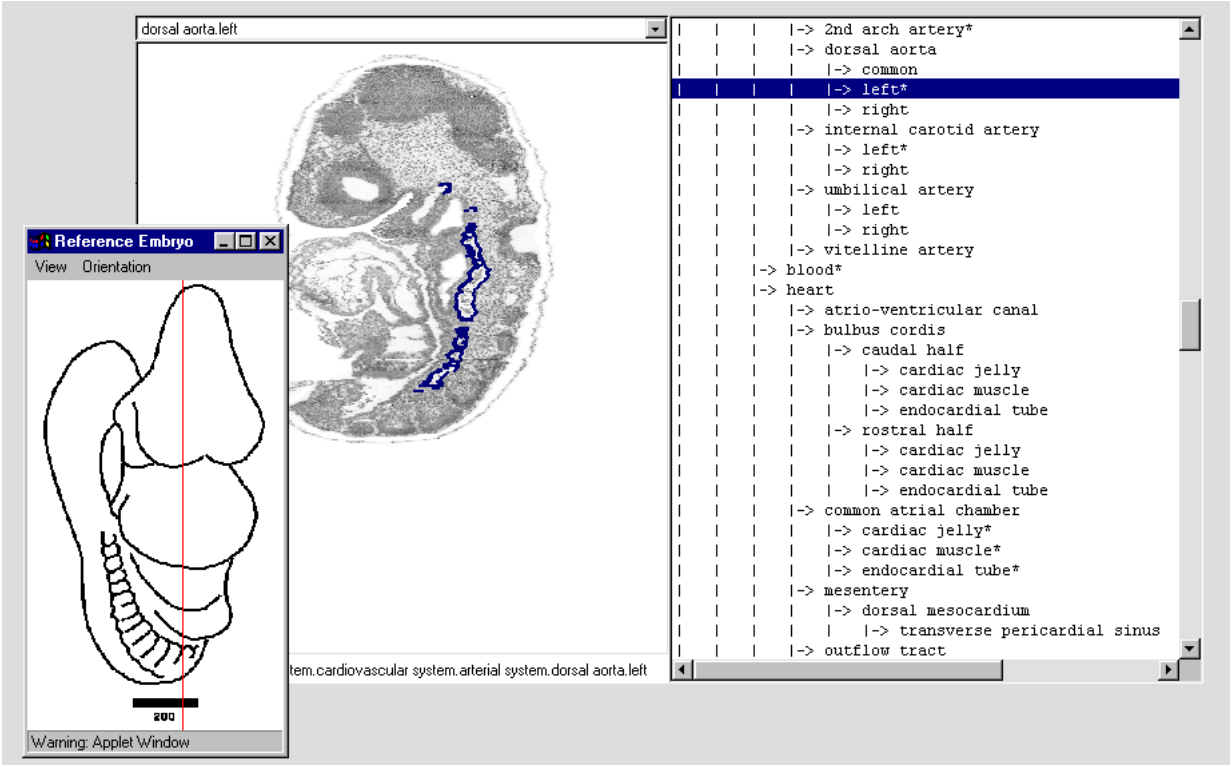

Figure 1: Screen shot of Mouse Atlas interface, displaying a Theiler Stage 14 embryo.

1056 matches.

The current project aims to provide a *robust free text interface* that will avoid these and other problems, (1) reducing the effort that a user needs to expend in finding anatomical structures of interest; (2) better organising search results; (3) improving *recall* by reducing the number of times that no match is found; and (4) improving *precision* over that which is possible using substring matching on individual component names. To do this, we have been extracting from the *Mouse Anatomical Nomenclature* another resource comprising a set of Natural Language (NL) phrases that uniquely denote the parts of the embryo. To expand this set, we have semi-automatically culled other anatomical specifiers from a textbook on developmental anatomy made available to us in electronic form. The interesting challenges this has posed are described in the remainder of the paper.

3 Reducing User Effort

As already noted, *component terms* in the Mouse Anatomical Nomenclature (as in those for other model organisms) are not unique designators for anatomical structures: the only unique designators are *path specifications*. Thus technically, the only way a user can uniquely select an anatomical structure of interest is to enter the entire path name (or to find it through navigating an embryo stage tree down from its root).

However, there are developmentally valid notions of uniqueness with respect to which some of the 1416 component terms associated with the 13737 nodes in the 26 Theiler Stage trees of the *Mouse Anatomical Nomenclature* can be taken to be unique.[4]

The first such notion of uniqueness can be associated with an anatomical structure that develops by some Theiler Stage j and then persists under the same name through subsequent stages. In the *Mouse Anatomical Nomenclature*, this situation corresponds, first off, to path specifications that differ only in their root note (which designates the embryo

[4]The size of Theiler Stage trees ranges from 3 nodes in stage 2 (early development) to 1739 nodes in stage 26 (pre-birth), with the average size being 528 nodes.

at the corresponding Theiler Stage). If the component term at the leaf does not occur elsewhere in the nomenclature outside this path specification, then this component name can be classified as unique.

This notion of uniqueness was found to hold of 1017 of the 1416 component terms, including CARDIOGENIC PLATE, CRANIUM, etc. This meant that approximately 11200 nodes were covered, with somewhat over 2500 still remaining. These 1017 component terms could *potentially* be used to access gene expression data associated with some or all of the Theiler Stages through which the uniquely designated structure exists, except for a problem that we will mention shortly.

The second notion of uniqueness is an extension of the first. Before anatomical structures are fully formed, they tend to be referred to by names that denote the same anatomical structure but also convey that it is not fully formed – for example, the FUTURE FOREBRAIN. Such component terms can be linked with the component term of the structure they develop into, treating the two together as a unique designator across the extended sequence of Theiler Stages. A user seeking gene expression data for the FOREBRAIN without specifying a particular Theiler Stage, could then select from stages 15-16, which contain the FUTURE FOREBRAIN, as well as from stages 17-26, which contain the FOREBRAIN.

In some cases, finding such terms is easy – specifically, when paths in adjacent stages differ only in their root node and their potentially co-designating leaf terms. In other cases, the process is complicated by the fact that their containing anatomical structures are themselves developing and changing. This creates additional differences in paths that should be taken to co-designate in this lineage sense. The difference may simply be in the particular component term associated with a non-terminal node – e.g. FUTURE FOREBRAIN is a child of FUTURE BRAIN in stages 15-16, while FOREBRAIN descends from BRAIN in stages 17-26. These cases can be identified by verifying that the intermediate structures are themselves in a lineage relation. But the paths may also differ in *length*, the earlier stage path being longer than its corresponding path in the next stage. This is because the earlier stage specifies the tissue from which the structure is developing from – for example, the fibula develops from the lower leg mesenchyme. So the path

EMBRYO.LIMB.HINDLIMB.LEG.LOWERLEG.
MESENCHYME.FIBULA

is a unique designator in Stage23, becoming

EMBRYO.LIMB.HINDLIMB.LEG.LOWERLEG.
FIBULA

in Stages 24 to 26. To recognise such cases, we need to analyse what nodes contribute to differences in path length and decide whether two component terms co-specify on that basis.

Again, when these patterns are encountered, the component names, providing they are not involved in any other initial tree paths, can be classified as being unique, further reducing the number of names to be disambiguated. So far 44 of these lineage patterns have been identified, further eliminating approximately 118 from the set of ambiguous component terms.

The third notion of uniqueness that can be used for identifying component terms that can serve as unique designators can be called *group uniqueness*. For example, although the component term TOOTH appears in different path specifications, whose corresponding internal nodes are not pairwise equivalent – e.g.

EMBRYO.ORGAN SYSTEM.VISCERAL
ORGAN.ALIMENTARY SYSTEM.ORAL
REGION.LOWER JAW.TOOTH

EMBRYO.ORGAN SYSTEM.VISCERAL
ORGAN.ALIMENTARY SYSTEM.ORAL
REGION.UPPER JAW.TOOTH

where LOWER JAW does not co-specify with UPPER JAW, the pairwise different nodes may correspond to structures whose anatomical/developmental properties can, in the context of gene expression, be considered the same. This notion of group uniqueness was found to hold of nine component terms, covering approximately 260 of the remaining nodes.

Before moving from individual component terms that turn out to be unique designators for anatomical structures, to short sequences of such terms, we need to point out a separate problem in actually using them in a user interface. The problem follows

from a design decision made in the development of these anatomical nomenclatures that supports their intended use by biologists as clear and succinct structural descriptions of an embryo. Specifically, while an individual component term may turn out to be a unique specifier with respect to the Nomenclature, outside the context of its tree path, it may not signify to a biologist what it is intended to. For example, while the term LOOP has been found to uniquely denote the same anatomical structure as

EMBRYO.ORGAN SYSTEM.VISCERAL ORGAN.
ALIMENTARY SYSTEM.GUT.MIDGUT.LOOP

a biologist would never simply use "loop" to refer to the the loop of the midgut. Similarly, while the term DISTAL uniquely designates the same anatomical structure as

EMBRYO.LIMB.FORELIMB.JOINT.
RADIO-ULNARJOINT.DISTAL

"distal" on its own is not how any biologist would refer to the joint of the radius and ulna bones that is furthest from the shoulder.

For the interface we are developing, we need to replace these albeit unique component terms with phrases that are more natural to use in specifying these structures.

Turning now to component terms that are not unique in any of the senses discussed so far, it still does not appear to be the case that a user need enter an *entire* path specification to refer to its associated anatomical structure. In many cases, a *sub-path* specification of two, or in some cases, three component terms appears sufficient to specify the anatomical structure of interest.

To find where shorter sub-paths would serve as a source of unique designators, we enumerated all sub-paths from nodes with a non-unique associated component term (either leaf or internal node) to the root of their corresponding Theiler Stage tree (i.e., paths being specified in child-parent order). This revealed many cases where a unique two- or three-component path specification would disambiguate an otherwise ambiguous term, and allowed us to cover another 156 component terms via the 2-component terms and 50 via the 3-component terms. This has left only 58 of the original 1416 component terms (and 1159 corresponding nodes in the Nomenclature out of the original 13737) for us to investigate other methods of finding unique designators for.

The question is what Natural Language phrases these multi-component terms correspond to, since it is such phrases that would be used in an interface, not sequences of component terms. Slight variations in what the parent-child relations correspond to mean there are three different phrasal patterns for two-component sub-paths: **(1)** In cases where the child and parent nodes are in a part-whole relation and both are realised as nouns – e.g., a child with component term CAPSULE descending from a parent LENS, or a parent CORTEX or a parent OVARY

EMBRYO.ORGANSYSTEM.SENSORYORGAN
.EYE.LENS.CAPSULE

EMBRYO.ORGANSYSTEM.VISCERAL ORGAN.
RENAL/URINARYSYSTEM.METANEPHROS.
EXCRETORY COMPONENT.CORTEX.
CAPSULE

EMBRYO.ORGAN SYSTEM.VISCERAL ORGAN.
REPRODUCTIVE SYSTEM.FEMALE.
OVARY.CAPSULE

the multi-component term can be realised as a phrase CHILD OF PARENT, generating the three uniquely specifying phrases "capsule of lens", "capsule of cortex" and "capsule of ovary". Alternatively, a natural phrase of the form PARENT CHILD, (i.e. "lens capsule", "cortex capsule" and "ovary capsule" can also be constructed as a natural way of describing the anatomical structure that the path denotes.

(2) In cases where the child and parent nodes are in a part-whole relation, but the component term associated with the child is an adjective such as LEFT, UPPER or ANTERIOR, then the pattern CHILD PARENT can be used to form an appropriate phrase. For example, the path specification

EMBRYO.ORGANSYSTEM.CARDIOVASCULAR
SYSTEM.VENOUSSYSTEM.VENACAVA.
INFERIOR

can be accessed by the phrase "inferior vena cava".

(3) In cases where the child and parent nodes are in a set-instance relation, as in the case of

EMBRYO.ORGAN SYSTEM.VISCERAL ORGAN.
ALIMENTARY SYSTEM.ORAL REGION.
GLAND.PITUITARY

again the pattern CHILD PARENT can be used to form an appropriate phrase – for example “pituitary gland” in this case. Phrases thus formed from multi-component sub-paths may again be unique with respect to an interval of Theiler Stages, or with respect to lineage within the Theiler Stages, or with respect to a group.

4 Improving the Display of Search Results

Currently, within the interface to the Gene Expression Database, one can search for an anatomical structure of interest within a single tree or across all stages. A query across all Theiler Stages (TS) results in a list of all stages with a matching component, and associated with each stage is a path specification terminating at a matching component name. This does not easily enable the user to locate the specific entity they are interested in. If the term is not unique, then the results contain all possible anatomical structures the query could represent. For example, a sub-string match on the phrase “lumen” results in

4 TS12 term(s) matching query “lumen”:

future spinal cord;neural tube;neural lumen
foregut diverticulum;lumen
hindgut diverticulum;lumen
midgut;lumen

5 TS13 term(s) matching query “lumen”:

future spinal cord;neural tube;neural lumen
foregut diverticulum;lumen
hindgut diverticulum;lumen
midgut;lumen
foregut-midgut junction;lumen

7 TS14 term(s) matching query “lumen”:

future spinal cord;neural tube;neural lumen (future spinal canal, spinal canal)
hindgut diverticulum;lumen
midgut;lumen
foregut-midgut junction;lumen
rest of foregut;lumen
foregut;pharyngeal region;lumen
otic pit;lumen

10 TS15 term(s) matching query “lumen”:

optic recess (lumen of optic stalk)
future spinal cord;neural tube;neural lumen (future spinal canal, spinal canal)
hindgut diverticulum;lumen
midgut;lumen
pharynx;lumen
foregut-midgut junction;lumen
hindgut;lumen
rest of foregut;lumen
foregut;oesophageal region;lumen
otic pit;lumen

Locating an entity of interest amongst all these tree paths can be an arduous task. Even if the term is unique, the same entry will be repeated across multiple stages, leading to a visual search problem.

An alternative, cleaner way of presenting search results is to take the matching component terms as the primary display key and associate it with a list of stages where its corresponding path specification occurs. For non-unique search queries such as the example above, this displays as

future spinal cord; neural tube; neural lumen:
Stages 12, 13, 14, 15, ...
foregut diverticulum; lumen:
Stages 12, 13
hindgut diverticulum; lumen:
Stages 12, 13, 14, 15, ...
midgut;lumen:
Stages 12, 13, 14, 15, ...
foregut-midgut junction; lumen:
Stages 13, 14, 15, ...
rest of foregut; lumen:
Stages 14, 15, ...
otic pit; lumen:
Stages 14, 15, ...
optic recess:
Stages 15, ...
pharynx; lumen:
Stages 15, ...
foregut; oesophageal region; lumen:
Stages 15, ...
foregut; pharyngeal region; lumen:
Stages 14

We will, of course, have to verify that this form of display better facilitates users finding the structure(s) and stage(s) of interest to them.

5 Increasing Precision

The introduction of phrases based on more than one component term within the Mouse Anatomy Nomenclature significantly reduces the number of

irrelevant matches compared to searches based on a single comonent term.

To continue our example with CAPSULE from Section 3, the current situation is that no results will be found if "cortex capsule" is entered as a search query. If the user then simply searches for "capsule" across all stages, 31 instances will be returned. However, although all sub-paths leading to CORTEX.CAPSULE are returned, the other 87% of the results are irrelevant to the user's intention. If the multi-component terms are included as uniquely designating replacements of existing terms, 100% recall would be maintained, while increasing precision to 100% percent.

As there is some systematicity involving parent and child component terms, these terms can be automatically generated. Within the nomenclature there are three typical patterns. These can be formalised as:

- the child being a descriptor of the parent e.g. superior vena cava
- the child being a part of the parent e.g. ovary capsule or capsule of ovary
- the child being a member of the parent set e.g. pituitary gland.

Of course, recognising which tree path belongs to which of the patterns above requires biological expertise, but once identified new terms can be generated that are more likely to be used naturally to refer to the relevant anatomical components. Once in place these new terms can increase recall, with high precision.

6 Increasing Recall

While the *Mouse Anatomical Nomenclature* was designed to specify every anatomical structure within the developing mouse embryo, it does not contain all the terms that developmental biologists might use to refer to anatomical entities. Although some synonyms have been explicitly recorded in the nomenclature, no attempt has been made to be exhaustive.

In order to increase the *recall* of user searches for anatomical structures, we have undertaken to increase the number and range of synonyms for elements of the nomenclature, by semi-automatically analysing texts likely to contain terms related to the developmental anatomy of the mouse.

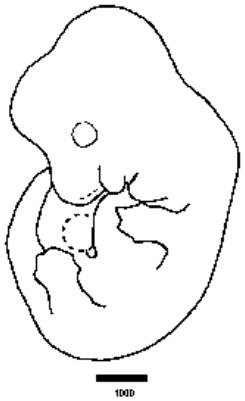

Figure 2: Schematic of a Theiler Stage 20 embryo

To demonstrate the potential value of this approach, we first manually reviewed the short textual description that accompanies each Theiler stage within the Mouse Atlas, highlighting the main features of the stage – for example, this text accompanies the schematic of TS20 (Figure 6):

> The handplate (anterior footplate) is no longer circular but develops angles which correspond to the future digits. The posterior footplate is also distinguishable from the lower part of the leg. It is possible to see the pigmentation of the pigmented layer of the retina through the transparent cornea. The tongue and brain vesicles are clearly visible.

We collected all the noun phrases (NPs) that could potentially refer to an anatomical structure and found, within the $<$1400 words comprising the descriptions, 25 anatomical terms that were not included in the nomenclature either as component terms or as synonyms for component terms. Since the same people wrote these textual descriptions as developed the Nomenclature, it shows how difficult it is to record all terms used for anatomical structures without systematic effort.

To support such a systematic effort, we have been applying basic text analysis software to a textbook on developmental anatomy (Kaufman and Bard, 1999), including a tokenizer, part-of-speech tagger and NP chunker, the latter two from

the Language Technology Group (LTG) in Edinburgh (http://www.ltg.ed.ac.uk), as well as additional scripts – in order to identify noun phrases, from which we then extract those most likely to refer to an anatomical structure. The latter have then been discussed with our domain expert, Davidson.

Because neither the POS tagger nor the chunker were specially trained for this type of technical text, their performance was rather weak. Chunker output from the chapter on the heart, produced 5547 phrases, of which 2465 were considered to be NPs. 2.4% (i.e. 74) of the 3082 claimed non-NPs were obvious *false negatives* involving the terms *venacava*, *septum primum/secundum* and *ductus arteriosus/venous*. Of the terms classified as NPs, 3.7% (i.e. 92) were found to be *false positives*. Most of these errors involved words that could be classed as verbs or nouns adjacent to true NPs, regarded as plural nouns but which, in context, were acting as verbs, e.g. *the ostium secundum forms*.

From the true NPs, we removed pronouns, numbers, authors names, and plural terms whose singular was also present. This left 451 NPs, of which 82 were found to be exact matches for component terms and eight, for the synonyms in the Nomenclature. These were also removed, leaving 361 possible anatomical terms not found in the Nomenclature.

We then used a common technique to reduce this set by only considering NPs headed by or modified by a frequent head or modifier from within the set of component terms (Bodenreider et al., 2002). Here, frequent meant $\geq$3 times. For example, CAROTID, FIBROUS and ENDOCARDIAL are frequent modifiers, while ARTERY, SEPTUM and TISSUE are frequent head nouns. Of the 361 remaining NPs from the heart chapter, 115 shared a high frequency head noun with terms already in the Nomenclature, while 105 shared a high frequency modifier with terms in the Nomenclature. We considered these 220 NPs *probable* anatomical terms, with the remaining 141 being *possible* anatomical terms. These two sets are now being reviewed to identify which are synonyms for anatomical structures identified in the Nomenclature, which denote structures or groups of structures that have not be recorded in the Nomenclature, which are reduced co-referring NPs, and which are not anatomical terms after all.

7 Future Work

One result of this work has been to catch both structural and terminological inconsistencies in the *Mouse Anatomical Nomenclature* because our extractions allow biologists to easily see differences between one branch and another or between one tree and another in the Nomenclature.

With respect to an enhanced interface to the gene expression data, we are now ready to take the results of our analyses and use them to provide a potentially more effective way of searching for relevant anatomical structures and displaying the results. We have a method to mine for additional synonyms for anatomical terms, that we will apply to additional texts, ideally after re-training the POS-tagger and chunker to better reflect the types of texts we are dealing with.

Similar Nomenclatures of developmental anatomy exist for other model organisms, including drosophila, zebrafish and human. These too will be used to index gene expression data for these organisms, eventually supporting cross-species comparison of gene expression patterns and further understanding of development. So we believe that developmental anatomy provides a rich domain in which to apply (and learn to extend) Natural Language tools and techniques.

Acknowledgements

The authors would like to thank Dr. Jonathan Bard for supplying us with electronic versions of two chapters from *The Anatomical Basis of Mouse Development* and answering many of our questions.

References

Olivier Bodenreider, Thomas Rindflesch, and Anita Burgun. 2002. Unsupervised, corpus-based method for extending a biomedical terminology. In *Proceedings of the ACL Workshop on Biomedical Language*, Philadelphia PA.

Matthew H. Kaufman and Jonathan Bard. 1999. *The anatomical basis of mouse development*. Academic Press.