Institute for Computer Graphics and Vision
Graz University of Technology
Graz

Metabolic Pathways Influencing Gene-Expression Analysis
Thesis paper

Marc Streit
mstreit@icg.tugraz.at

May 2007

Supervision:
Univ. Prof. DI Dr. techn. Dieter Schmalstieg
DI Micheal Kalkusch

Abstract

Ich versichere hiermit, diese Arbeit selbständig verfaßt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und mich auch sonst keiner unerlaubten Hilfsmittel bedient zu haben.

I hereby certify that the work presented in this diploma thesis is my own and that work performed by others is appropriately cited.

Acknowledgements

Contents

Chapter 1

Introduction

We focus on investigation of existing pathways and not the creation of new ones (references?).

Chapter 2

Problem definiton

[Interesting term: post-genomics - paper: Tracking the Shift to postgenomics]

Tool that supports biologists and pathologists to understand the results of their experiments.

[Describe data increase rate - specially in context of high-throughput methods]

Chapter 3

Related work

In this chapter first pathways are described in general followed by a detailed consideration of pathway visualization and current tools and frameworks. Next the term gene expression is introduced and again state-of-the art framework are described roughly. [Missing introductory sentence for information visualization]. The combination of these two ermerging research areas leads to gene-expression application on biochemical pathways. Finally related work in the field of information visualization that is needed in respect of the former described fields is discussed.

3.1 Metabolic Pathway Visualization

3.1.1 Medical Pathways

[Metabolism is defined in the book Metabolism at a glance and is cited in (BALJ06)]

Metabolic networks are defined in (BALJ06) as interconnected metabolic pathways. [Write some more sentences about metabolic networks and why they are important to be investigated globally.] Figure figure 3.1 shows a whole snapshot of a metabolic pathway. The fact that these posters of the global metabolic network are still mounted on the wall in numerous laboratories illustrates why alternative ways of visualization needs to be explored.

3.1.1.1 Definition

Metabolic pathways aim on modelling cellular functions in graphs. Basic building blocks of pathways are chemical compounds (so called metabolites) and enzymes. The chemical compounds act as substrates and products of chemical reactions. Enzymes catalyze these reactions by using substrates as input. The output of the chemical reaction is a product which is in turn a compound. This process is described in LINK.

Pathways can be divided in metabolic pathways and regulatory pathways such as signal transduction cycles. The former are in the scope of that paper because metabolism can be modeled in reference pathways which are valid for several organisms(KG00). Outgoing from that reference map KEGG can generate organism specific pathways. On the other hand regulatory pathways vary in detail for specific organism and it is much harder to find a common reference pathway(KG00).

[Describe additional information as co-factors etc.]
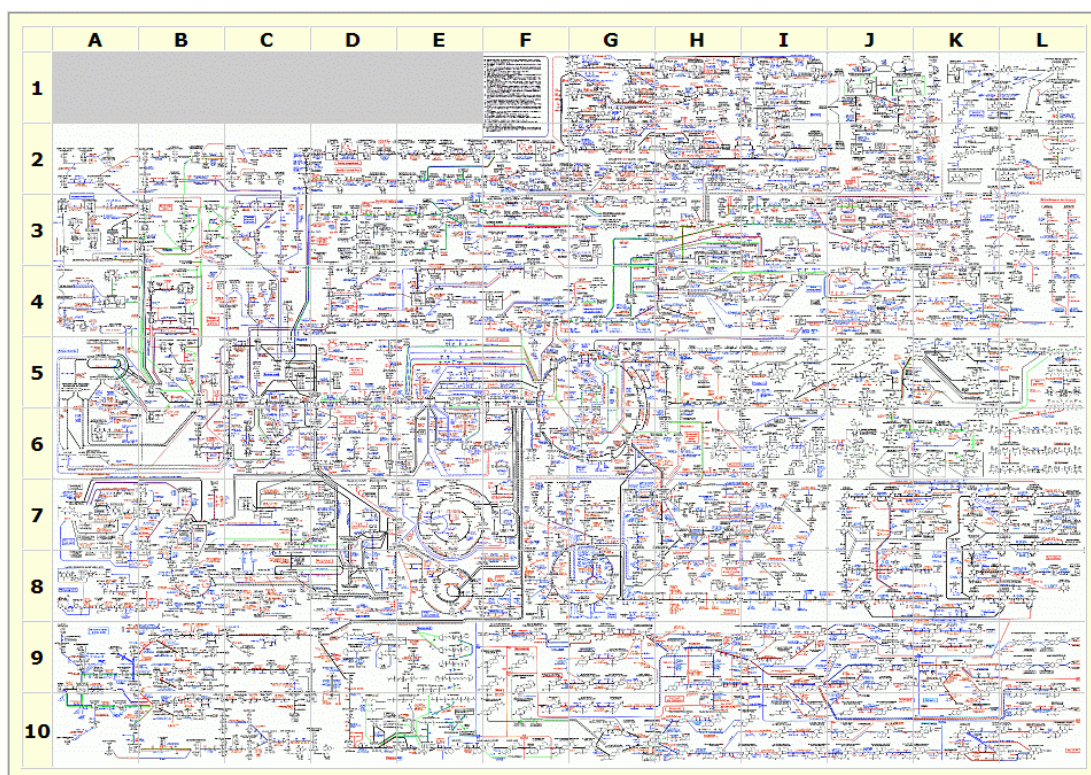
Figure 3.1: *Roche Applied Science wall chart on metabolic pathways TODO Insert Reference*

3.1.1.2 Layouting and graph drawing algorithms

[Defintion of metabolic networks regarding graph theory is presented in (BALJ06).] A very good explaination of metabolic networks and pathways in respect of graph theory is given by Bourqui et al.(BALJ06). The metabolic network is a bipartite graph where the set of nodes is divided in two disjoint subsets of nodes whereas no nodes inside a group are connected. In the case of metabolism that means that enzymes are not directly connected to other enzymes. The same applies for chemical compounds. Metabolic pathways are partially cyclic graphs. Directed and undirected edges are present. Nodes can occure in multiple times in the same graph as well as in foreign graphs.

Determined by historical development metabolic networks are often handmade. I.e. experts incorporate their meta-knowledge in the specific domain to position nodes and route edges. Still this approach is widely spread in the community due to the complexity of the graphs. Graph drawing algorithms are needed for automatic generation of pathways. Common graph drawing algorithms do not provide sufficient results in respect of metabolic pathways because they focus on different constraints(BR01). Becker et al. state that common graph layouting alogrithms are optimized to fullfil the following cirterias:

- Planarity
- Minimal edge crossings

- Minimal drawing area

- Maximal symmetry

In matters of metabolic graphs it is hard to define a simple set of criterias and constraints because the complexity and diversity. Therfore special algorithms need to be developed that are capable of handling the biochemical properties.

One of the first attempts to dynamically model metabolic networks was done by Karp et al. starting in 1993(KR93; KP94a; KP94b). As metabolic network became more and more diversified and complex over time also graph drawing approaches needed to be adapted and enhanced. Becker et al. proposed an algorithm that builds up on the ideas of Kerp et al. and enhances them by including the topological structure. As a starting point the method seeks for the longest cycle inside the graph. The rest of the nodes are classified as inner and outer components which are positioned using a spring embedded algorithm.

In the recently published work by (BALJ06) a graph drawing algorithm is published that addresses the visual identification of reaction cascades among distinct metabolic pathways by forming a single concatenated graph [should I put a figure here?].

3.1.1.3 2.5D and 3D pathway visualization

3.1.1.4 Pathways visualization using Virtual Reality (VR)

[Insert short definition of virtual reality - Judith bringt Referenz.]

In 2002 Rojdestvenski et al. ((RC02; Roj03)) proposed a tool called Metabolic Network Visualizer (MNV) for the usage of virtual reality in context of metabolic pathways by using Virtual Reality Modeling Language (VRML). They also discuss advantages and possible problems when using a real 3D representation of the biomedical network. Finally a hierarchical postitioning of the graphs in 3D space as flat 2D graphs is proposed to keep the 2D graphs with that the biologists are familiar with. (DJ03):[paper contains a good definition of metabolic pathways as graphs] In (YWCND06) the MetNetVR is presented which focuses on hierarchical relationships in pathways (-> also in (DGC$^+$04)). The root node of the graph could be the metabolic network. Childs of the network are several pathways which subnodes are their molecules. For exploring the information space detail-on-demand techique (-> link to other section) is applied. Using a tracked input device orientation and position for ray-picking in an virtual environment the user can expand a single node pathway representation to show all contained molecules.

3.1.2 State-of-the-art Frameworks

Brandes et al. propesed a 2.5D method to visualize differences accross pathways for different spzecies(BDS04). This method enables the user to see evolutionary developments and relations amoung different species.

3.2 Gene Expression Visualization

3.2.1 Gene-Expression Analysis

Describe the Genome. Human genome is sequenced since year by name -> Reference Describe how genes code enzymes (special proteins).

Hierarchical clustering(SS02)

3.2.2 State-of-the-art frameworks

http://www.genome.jp/kegg/expression/
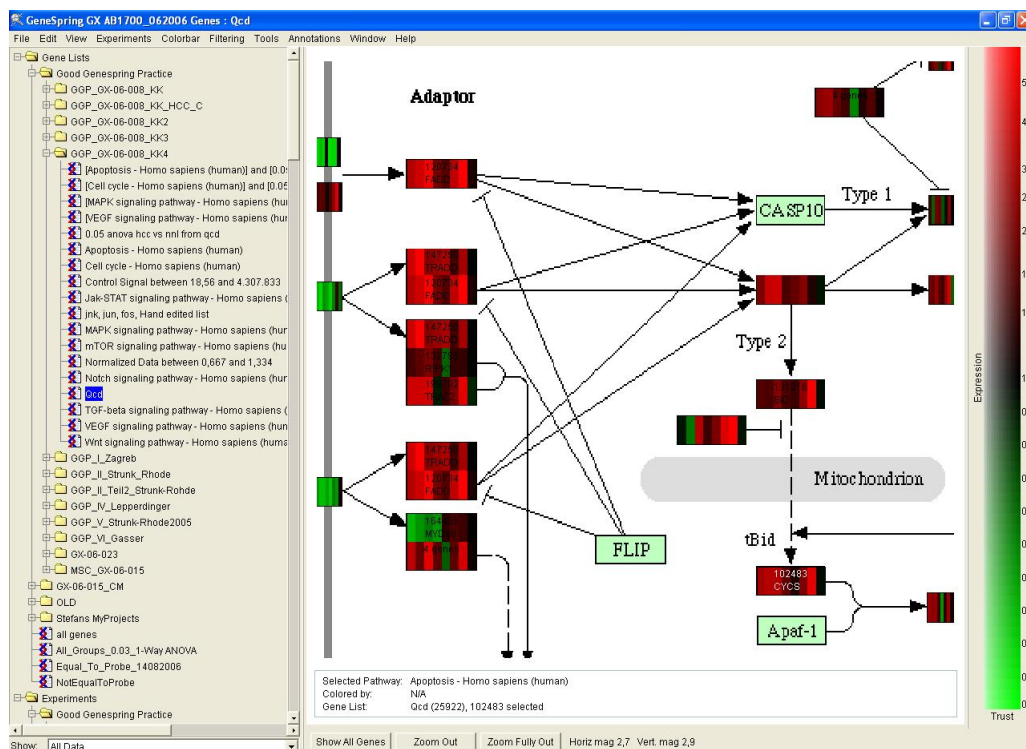
GeneSpring

Figure 3.2: *Screenshot of GeneSpring*

Panther

The PANTHER (Protein ANalysis THrough Evolutionary Relationships) classifies genes by functions[1]. [Reference?]

Pathway drawing tool called CellDesigner[2].

---

[1]http://www.pantherdb.org

[2]http://www.celldesigner.org
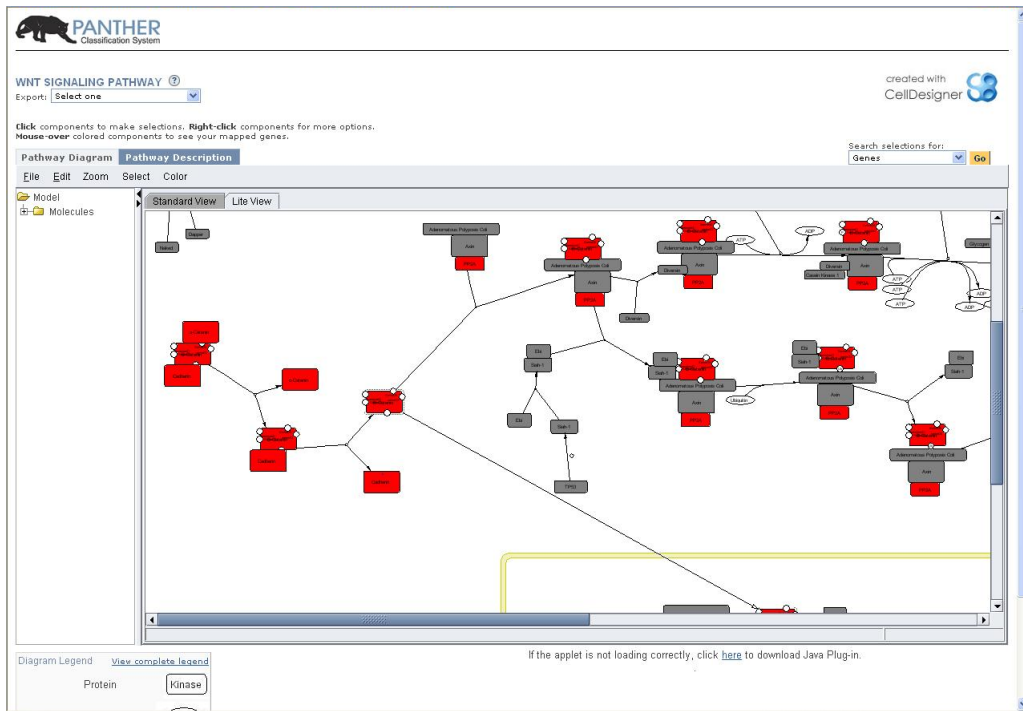
Figure 3.3: *Screenshot of CellDesigner*

3.3 Biomedical Databases

Criterias for successful databases are:

- Availability
- Update rate / update process
- ID Management (Unique Key)
- Links to other databases
- Software to access database (SOAP)
- Documentation
- Full data dump
- etc.

3.3.1 Gene centered databases

3.3.1.1 International Nucleotide Sequence Database Collaboration (INSDC)

The INSDC is a collaborative network of DNA sequence databases that consists of three major members:

- **GenBank:**
  GenBank is a part of the National Center for Biotechnology Information (NCBI)

which is intergrated in the United States National Library of Medicine (NLM). NLM belongs to the National Institutes of Health (NIH) that is self-described as the focal point for medical research in the United States.

- **European Molecular Biology Laboratory (EMBL):**
  EMBL

- **DNA Data Bank of Japan (DDBJ)**

INSERT table with current data sets in various databases Number of genes, proteins, enzymes.

3.3.1.2 Entrez

3.3.1.3 Gene Ontology (GO)

3.3.2 Pathway centered databases

3.3.2.1 Kyoto Encyclopedia of Genes and Genomes (KEGG)

The Kyoto Encyclopedia of Genes and Genomes (KEGG)[3] is a biomedical resource that started its online service in 1995 and belongs to the Japanese GenomeNet.

Provide 140 metabolic pathways for over ?? organisms.

KEGG pathways are represented by graphs which are layouted by hand and stored in static images (reference?). [Images are generated, how?] ok In the graph enzymes are visualized by rectangular nodes and compounds are modelled by small circles. Round rectangular nodes depict linked pathways. This circumstance points to the fact that a pathway is only an artificial subset of a huge complex network.

When a cellular function is valid throughout different organisms it is called strongly preserved. These general pathways are stored in the KEGG database as reference pathways. Each reference pathway can then be specialized for a specific organism. Figure BLA shows the methionine metabolism for home sapiens. Light green color coded enzymes depict proteins where the generating genes are known.

[Describe possiblities of KEGG in terms of database information and relation browsing - pathway coloring - hierarchical division of metabolic network]

Throughout the paper Methionine Metabolism serves as prime example. This special amino acid pathway is of average size and includes all types of nodes and edges as well as cyclic reaction cascades. Figure 3.4 shows the reference pathways of Methionine Metabolism while figure 3.5 depicts the specialized cellular function cycle for homo sapiens. The graphs share the same composition and connection of nodes but vary in the color coding of enzymes. Human genes that produce enzymes in the pathway are highlighted yellowish.

---

[3]http://www.genome.jp/kegg/
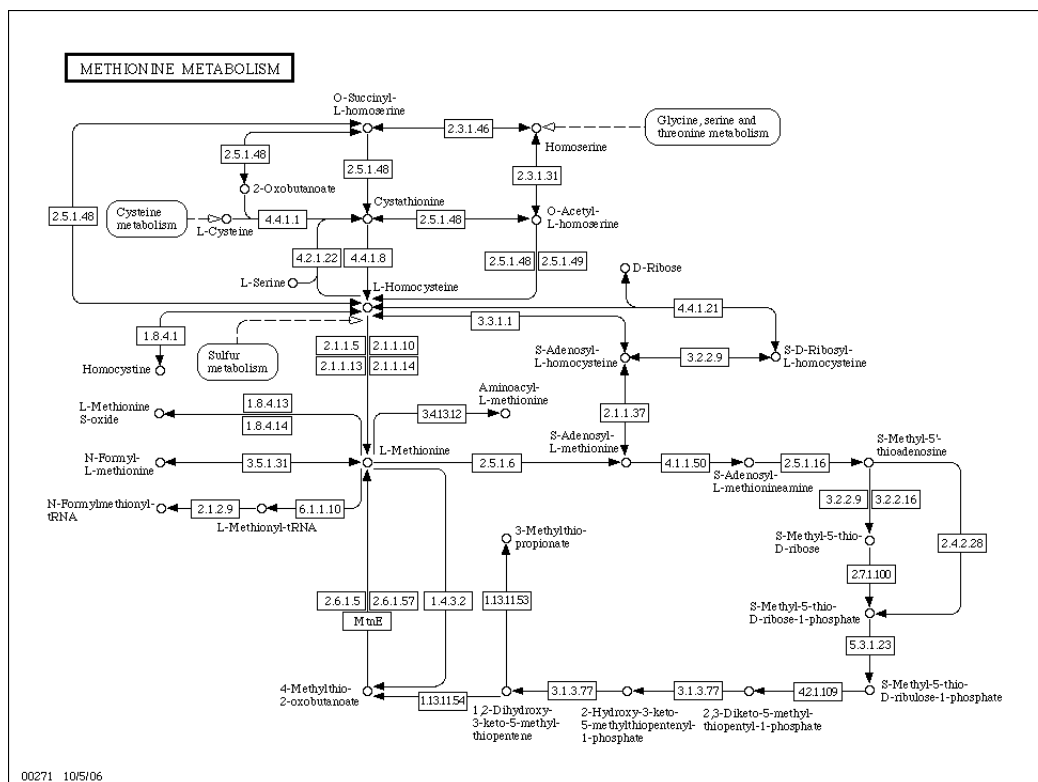
Figure 3.4: *KEGG methionine metabolism reference pathway*

3.3.2.2 BioCyc

3.3.2.3 ExPASy Molecular Biology Server

Reference: A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server (Appel et al.)

3.3.3 Biomedical Identification Numbers

During the last decade biochemical database projects sprang up like mushrooms. Some projects has proven their value and achieved to manifest themeselves inside the community. Each database project introduced their own identification numbers for genes, proteins, nucleotids, etc. The obvious problem is how to map these data and interconnect them among the misscellaneous databases. There were many attempts over the time to get this issue under control.

Enzyme ID

The Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) agreed on the Enzyme Commission number (EC number)[4](Bai00). Enzymes are divided by classes and a couple of subclasses which are seperated by a "."

---

[4]http://www.chem.qmul.ac.uk/iubmb/enzyme/

Figure 3.5: *KEGG methionine metabolism pathway for home sapiens*

(e.g.: 3.1.1.1). The Enzyme nomenclature database[5] is online provided by the Swiss-Prot platform. The release of 03-Apr-2007 contains 4026 active entries. Each entry contains:

- EC-number
- Official name
- Alternative names
- Reactions catalyzed
- Co-factors (if any)
- Cross-links to other databases

Gene ID

The nomenclature for enzymes become widely accepted however inn the case of genes unique identifiers are much more silvered. Several identifcation systems are still under usage but the main problem is the uniquity of the gene identifiers.

Pathway ID

3.4 Information Visualization Methods

(BSB03) (Spe07) (SM00)

3.4.1 2D vs. 3D

3.4.2 Multiple Views

3.4.3 Focus + Context

3.4.4 Linking & Brushing

3.4.5 Detail on demand

3.4.6 Semantic Zooom

3.5 Application of Gene-Expression Data onto Metabolic Pathways

3.5.1 State-of-the-art frameworks

PathwayExplorer [2005]

The PathwayExplorer is able to map genes on enzymes in pathways from various databases. Positive: Can handle all kind of identifiers (for example RefSeq, ...). Negative: Needs time to map genes on pathways. The result are static PNG images.
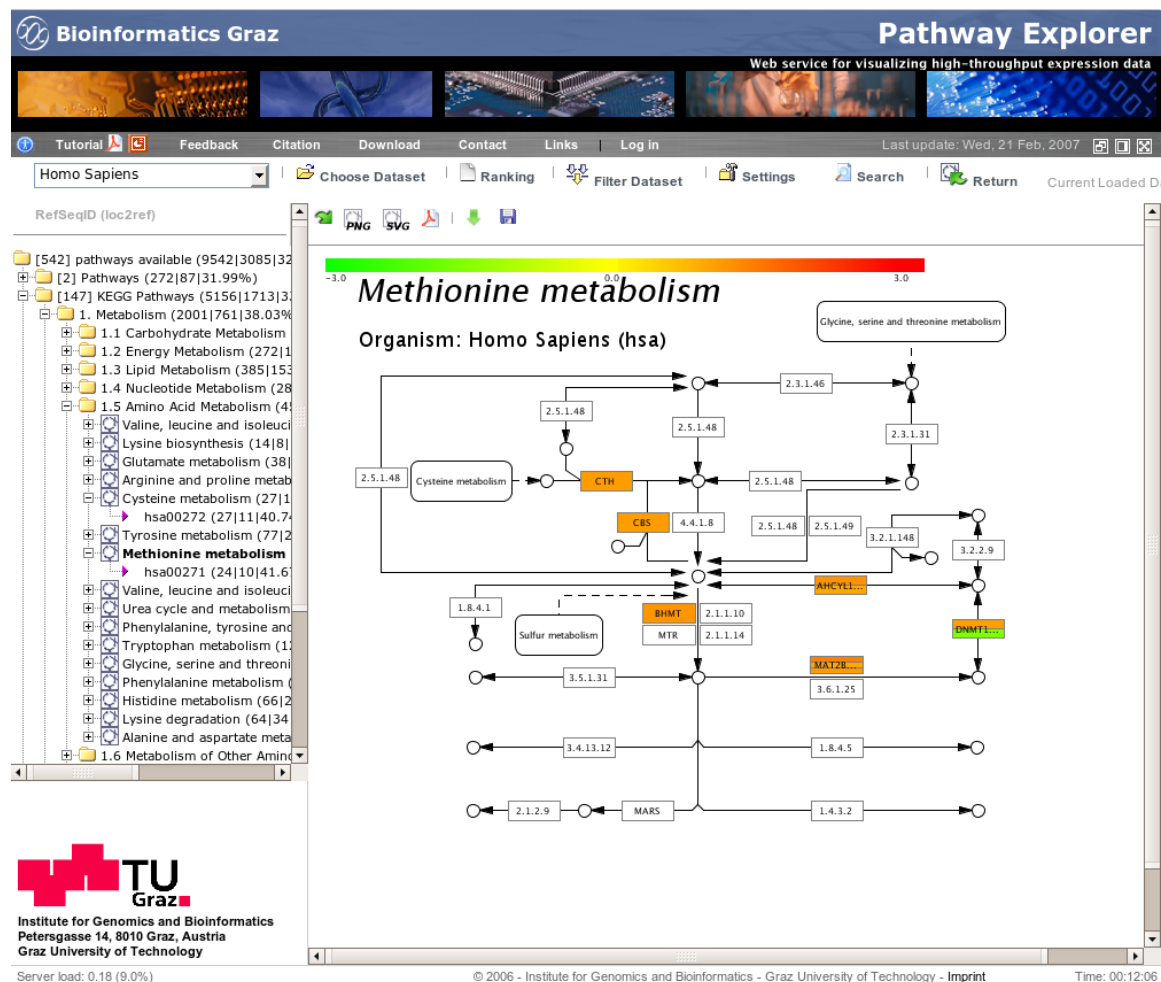
Figure 3.6: *PathwayExplorer*

Chapter 4

System architecture

The system design described in this chapter is part of the Cerberus visualization framework [Reference?]. First of all the module design of the framework is depicted. This design overview is followed by a rather detailed presentation of selected sub-modules. Finally the mapping process of the gene-expression data onto the metabolic pathways is conceptually outlined.

4.1 Overall Design

Platform independent

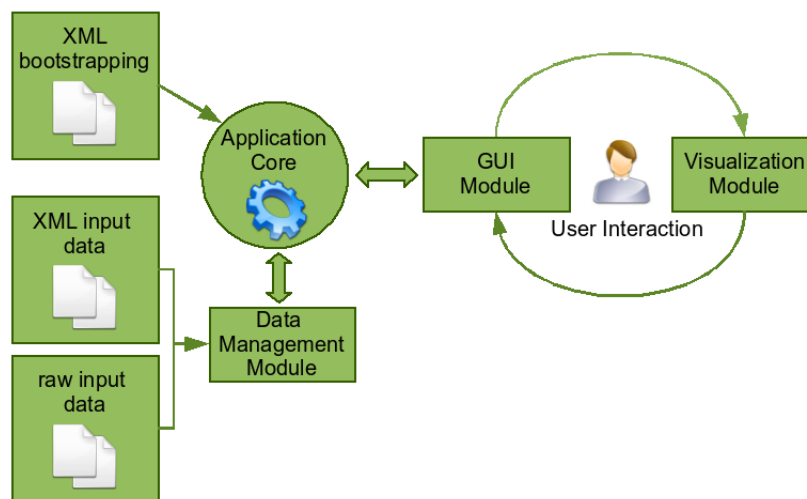Figure 4.1: *Application block chart*

4.2 Pathway Data Management

From information point of view metablic pathways are graphs. The nodes are enzymes and compounds. The edges can be differenciated in relations and reactions.

[The problem description why this solution is needed is very good described in (BALJ06).]

[Describe idea of data proxy]

[Insert UML diagram of pathway data management]
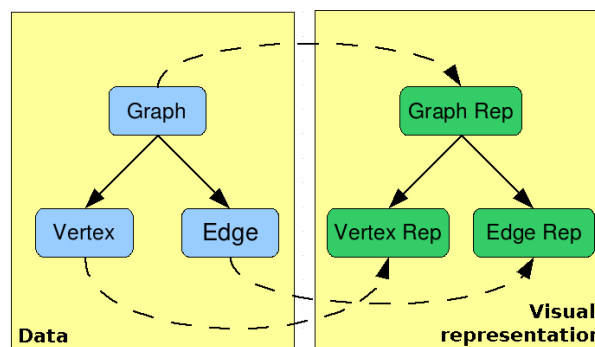
Figure 4.2: Model abstraction into data and visual representations
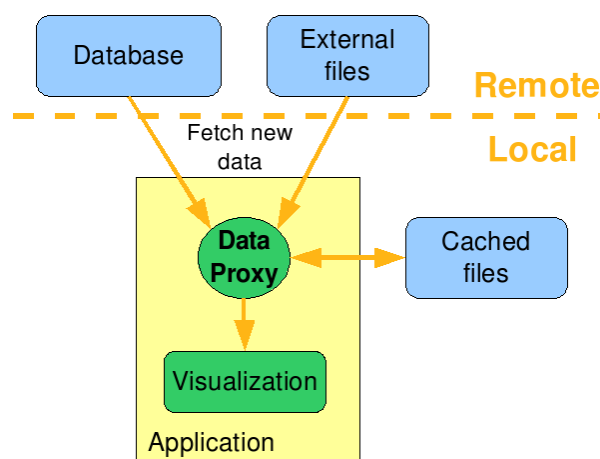
Figure 4.3: Data block diagram showing the data loading process

4.3 Graphical User Interface (GUI)

[Insert UML diagram of GUI-Manager & View-Manager] [Describe layouting of framework - container can hold several views]

[Define what we mean by a VIEW]

Views: - 2D Pathway - 3D Pathway - Scatterplot - Heatmap etc.

4.4 Data update (notification) mechanism

A data update is the process of changing data and in turn informing related parties. Objects of the systems are qualified to receive update notifications if they somehow act on this data. The receiver can decide what actions will be triggered.

[Insert UML diagram]

4.5 Enzyme-Gene Mapping

For realizing the mapping of enzymes onto the pathway we had to face the question how to store the data. Considering the fact that the framework has to handle several thousands of enzymes and up to the factor of 10 more genes the design descision needed to be well considered. The complete mapping tables are loaded into RAM at startup and are held during runtime. Of course, this approach increases the working memory consumption of our application considerable. However, the considerably high working memory consumption is consciously accepted because the requirements focuses cleary onb short mapping times.

For each external identification number (e.g. enzyme codes, gene accessions, etc.) that is fed into the system an internal ID is generated. The idea behind that procedure is that in every stage of the system the type of data can be fastly determined by investigating the ID. The disadvantage that this approach brings along is that an external to internal ID mapping table needs to be accessible at any time.

The mapping mechanism must take care of the fact that a bidirectional conversion of most ID types is required.
Example: User selects an enzyme inside a pathway and want to know which genes produce that particular protein (i.e. mapping from enzymes to genes). On the other hand in for example a scatterplot that visualizes gene-expression results the user might be interested in the encoding of which enzymes a particular gene is involved (mapping from genes to enzymes).
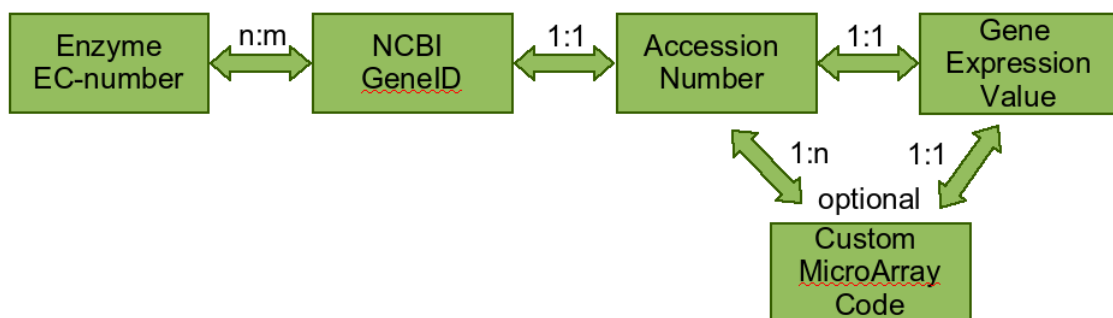
Figure 4.4: *Enzyme-Gene Mapping*

Our framework is capable of visualizing different gene mapping methods:

- Multiple gene mapping
- Time series gene-expression visualization

Map at end of pipeline points to storage (by holding the index as value)
GenomeId manager configuration via XML file

[Decide where to put the text] When multiple views are connected and the sent data needs to be mapped this procedure takes place inside the update callback method of the receiver. This policy origins in the consideration that only the target view knows how to map the

data because that is the point where the information is going to be visualized. This is also the perfect place for filtering the data (e.g. ignore data that is useless for the visualization in that view).

Chapter 5

Implementation

5.1 Used Technologies

5.1.1 Standard Widget Toolkit (SWT)

SWT[1] is an open source library for creating graphical user interfaces. The project was initiated 2001 by IBM under within the scope of the Eclipse framework[2]. In contrast to the standard Java toolkits like the Abstract Widget Toolkit (AWT) and Swing took the SWT designer an approach which is closer to the native operating system widgets. This results in a better performance and a native look of the widgets.

5.1.2 JOGL

Java OpenGL library[3] [Compare to Java3D]

5.1.3 JGraph

5.2 Data Loading

KEGG provides the data the XML (Extensible Markup Language) file format which made the job easy to parse the data in Java. We had to face the choice between between several ways of parsing offered by the Java Standard API. Most relevant methods are(NK02):

- **Simple API for XML (SAX):**
  TODO: Describe SAX here

- **Document Object Model (DOM):**
  TODO: Describe DOM here

KEGG pathway are defined in the KEGG Markup Language (KGML)[4]. KGML mixes information about the map structure and the visual representation (like colors, shapes, etc.). Here is a snip of the example pathways Methionine Metabolism in KGML stored in a XML file:

---

[1]http://www.eclipse.org/swt/
[2]http://www.eclipse.org
[3]https://jogl.dev.java.net/
[4]http://www.genome.jp/kegg/docs/xml/

```
<pathway name="path:map00271" org="map" number="00271"
title="Methionine metabolism"
image="http://www.genome.jp/kegg/pathway/map/map00271.gif"
link="http://www.genome.jp/dbget-bin/show_pathway?map00271">

  <entry id="1" name="ec:2.6.1.-" type="enzyme" reaction="rn:R07396"
  link="http://www.genome.jp/dbget-bin/www_bget?enzyme+2.6.1.-">
    <graphics name="2.6.1.-" fgcolor="#000000" bgcolor="#FFFFFF" type="rectangle"
    x="335" y="571" width="45" height="17"/>
  </entry>

  <entry id="2" name="cpd:C08276" type="compound"
  link="http://www.genome.jp/dbget-bin/www_bget?compound+C08276">
    <graphics name="C08276" fgcolor="#000000" bgcolor="#FFFFFF" type="circle"
    x="466" y="494" width="8" height="8"/>
  </entry>
  ...
  <reaction name="rn:R01402" type="irreversible">
    <substrate name="cpd:C00170"/>
    <product name="cpd:C04188"/>
  </reaction>
  ...
</pathway>
```

The alternative would habe been the usage of the KEGG application programming interface (API). The in Java implemented API employs the Simple Object Access Protocol (SOAP) for the transfer of XML information over HTTP. Due to the fact that real-time behaviour is in the main focus of the framework the SOAP alternative was too slow.

Problem of double data keeping.

5.3 Visualization Techniques

In the upcoming section used visualization methods are described and their realization is going to be discussed. Basically the section is divided in 2D pathway implementation and the 3D OpenGL pathways.

5.3.1 2D Pathway Implementation

The context information where in the hierarchical KEGG levels the presented piece of information is located would lead to a better orientation inside the metabolic network (JM03). Therefore this feature can be seen as a valuable extension for the future.
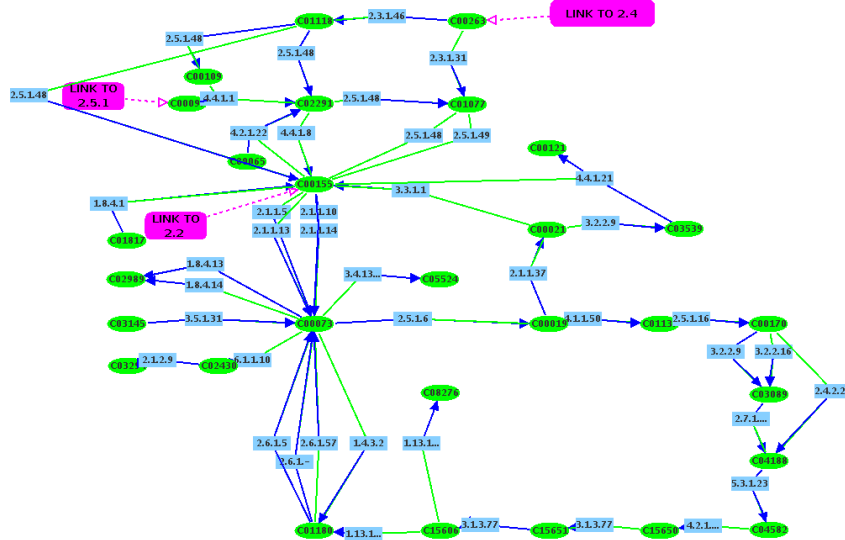
Figure 5.1: Sample pathway without any layouting modification.

5.3.1.1 Pathway Switching

Pathways are linked to related pathways. In the KEGG drawing style linked maps are represented by roundrectangular nodes inside pathway graphs. Therefore it is important to support fast switching between them. We implementing the switching action by replacing the currently shown graph which is triggered by a mouse event. This mechanism enables the user to freely navigate inside the metabolic network.

5.3.1.2 Hierarchical Pathways

KEGG added two abstraction levels above the metabolic pathways. The pathways are categorized in 10 groups which represent the highest layer. The middle level... Figure 5.3 shows the three KEGG abstraction levels. In the webbased KEGG version each abstract level has an imagemap in the back. In the imagemap rectangular portions are defined and provided with a hyperlink to the next layer below. We use the same mechanism for linking pathways between abstraction levels. In contrast to KEGG we store the imagemap information in XML which is parsed at bootstrapping. During runtime when a mouse event is triggered on the 2D pathway the clicked position is checked if it is contained in a predefined area. If a positive match is found the link for this area is returned and the requested next layer map can be loaded.

5.3.1.3 Neighborhood Visualization

In graphs where the layout aims on the positioning of nodes and edges to circumvent inter-sections neighborhood visualization occupies an important role. The geographic position of nodes is independent from their relational distance. Hence far away positioned nodes can be directly related to a node in focus whereas near nodes are e.g. several indirections
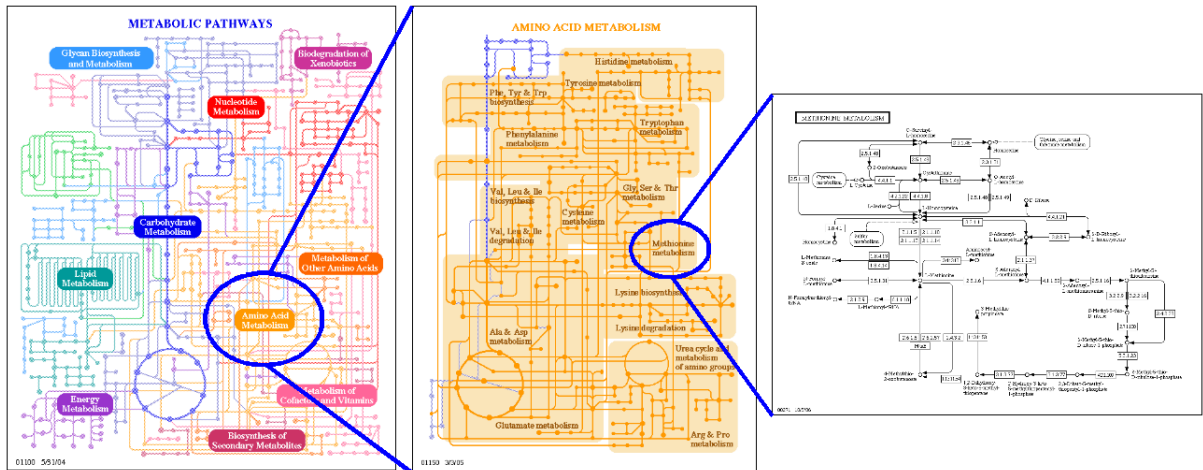
Figure 5.2: Sample pathway with background texture overlay (FIX screeny: labels are missing!).

away.

We implemented the neighborhood visualization using a modified Breadth-first-search (BFS) approach(CLRS01). The BFS algorithm visits all direct neighbors. After that procedure for each neighbor the same strategy is applied. If an already visited node is inspected it will be ignored. This assures that only the shortest path to a node is considered (for unweighted graphs). The algorithm is applicable on undirected and directed graphs. The alogrithm can be executed for arbitrary depth. Neighborhood coloring up to the depth of 3 seems to be the limit. The application of farer away neighborhood distances entails confusion at user's side because of the cyclic charakter of the graphs. The neighborhood algorithm for a distance of 3 is applied on the sample Methionine Metabolism pathway in figure 5.4. The selected node is colored red. The neighboring nodes are shaded from orange (distance 1) to light yellow (distance 3).

5.3.2 3D OpenGL Pathways

5.3.2.1 Hierarchical Display Lists

5.3.2.2 Pathway Texture Overlay

5.3.2.3 Showcase layouts

Caching the model-view matrix per pathway allows a free aranchement of the planar graphs in 3D space. This approaches enabled us to implemented two simple exemplary setups.

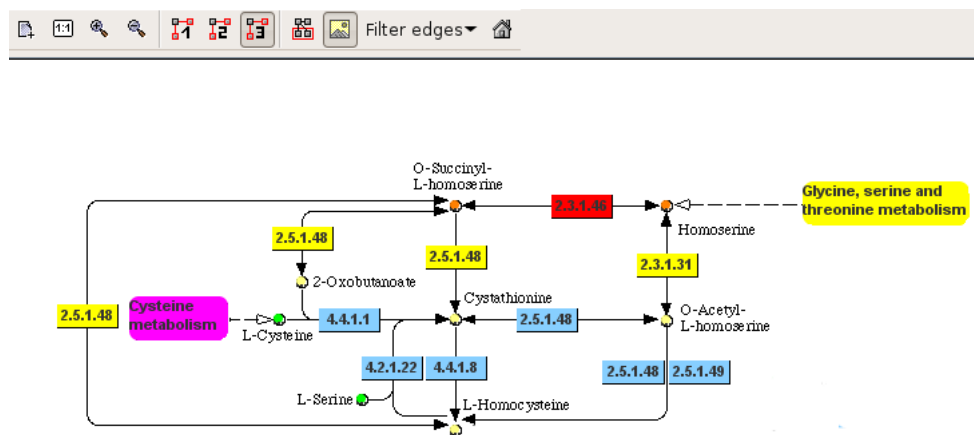Figure 5.3: *KEGG abstraction chain*

Figure 5.4: *Distance 3 neighborhood visualization*

Layered pathway setup

[Insert screenshot of a layered view]

Planar pathway setup

One graph in the center of user interaction. Metapher of a gallery.

[Insert screenshot]

5.3.2.4 Pathway Linking and Switching

5.3.2.5 Pathway Element Picking

Picking is the operation of selecting a objects in the scene. In our case we employ the picking action to select nodes inside the graphs. In OpenGL element picking can be achieved by different methods. The straight forward approach uses the OpenGL build in selection mode(SWND05). The user starts the picking operation by triggering a mouse-click event.
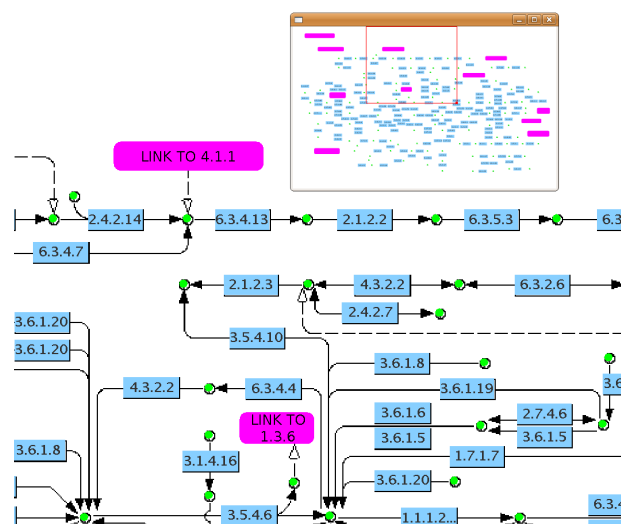
Figure 5.5: Portion of a sample pathway that can be interactively selected in the overview map.

The selection mode is entered and an area around the clicked x,y screen-coordinates is defined by appliying a picking matrix on the current matrix stack. Accordingly this region is then used by the selection render mode to restrict the rendered area. In selection mode the render method returns a hit record which contains the information about the intersecting objects inside the region. The size of the region determines the sensitivity of the selection.

In addition to the described method we used hierarchical object picking. In that case multiple names are returned for each hit. Applied to our use case the scene is made-up of several pathway graphs which represent the top-level elements in the hierarchy. In turn vertices inside the graph are on the second level. Hence the picking result gives us exact information which node in which pathway the user is interested in.

Increasing the number of visualized pathways in one scene also raises the probability of multiple objects under the curser. Therefore a picking action returns a set of hits. As each hit comes with a depth value the one which is nearest to the viewport is taken.

5.3.2.6 Enzyme-Gene Mapping

Multi hash maps (describe)
Templated hash maps parametrized by data type
For time series experiment the index remains the same - just the pointer to the storage is replaced
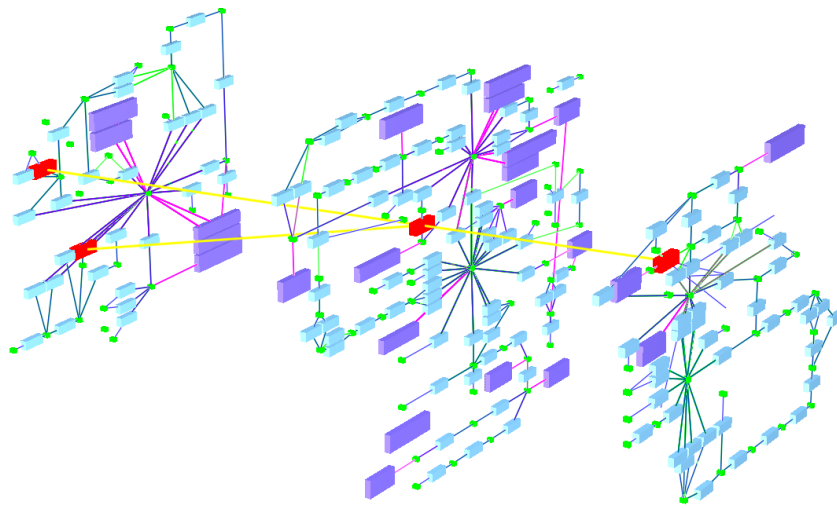
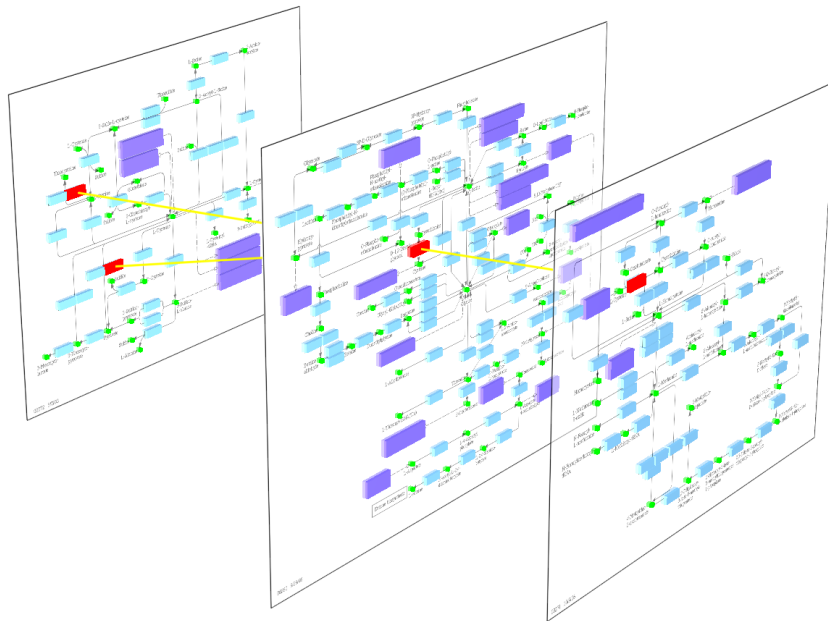Figure 5.6: 2.5-D OpenGL layered pathway with a user selection

Figure 5.7: 2.5-D OpenGL layered pathway with a blended texture overlay. The node in the middle layer is picked by a mouse-click event. Identical nodes in dependent graphs are interactively highlighted and connected to the selected one.

Chapter 6

Results

[Reflections from Zatloukal]

Chapter 7

Conclusion and future work

List of Figures

List of Tables

Bibliography

[Agr06]   AGRAWALA, Maneesh: Software Design Patterns for Information Visual-
          ization. In: *IEEE Transactions on Visualization and Computer Graph-
          ics* 12 (2006), Nr. 5, 853–860. `http://dx.doi.org/http://dx.doi.`
          `org/10.1109/TVCG.2006.178`. – Student Member-Jeffrey Heer. – DOI
          http://dx.doi.org/10.1109/TVCG.2006.178. – ISSN 1077–2626

[Bai00]   BAIROCH, Amos: The ENZYME database in 2000. In: *Nucleic Acids
          Research* 28 (2000), Nr. 1, S. 304–305

[BALJ06] BOURQUI, Romain ; AUBER, David ; LACROIX, Vincent ; JOURDAN, Fa-
          bien: Metabolic network visualization using constraint planar graph draw-
          ing algorithm. In: *IV '06: Proceedings of the conference on Information
          Visualization.* IEEE Computer Society. – ISBN 0–7695–2602–0, 489–496

[BDS04]   BRANDES, Ulrik ; DWYER, Tim ; SCHREIBER, Falk: Visualizing Related
          Metabolic Pathways in Two and a Half Dimensions (Long Paper). In: LI-
          OTTA, Giuseppe (Editor): *Graph Drawing, Perugia, 2003*, Springer, 111-122

[BFP+04] BRANDENBURG, Franz J. ; FORSTER, Michael ; PICK, Andreas ; RAITNER,
          Marcus ; SCHREIBER, Falk: BioPath - Exploration and Visualization of
          Biochemical Pathways. Version: 2004. `https://bscw.icg.tu-graz.ac.at/`
          `bscw/bscw.cgi/d27690/Springer_2004_BioPath%20-%20Exploration%`
          `20and%20Visualization%20of%20Biochemical%20Pathways.pdf`.   In:
          J?ï>½NGER, Michael (Editor) ; MUTZEL, Petra (Editor): *Graph Drawing
          Software.* Springer, 215 – 236

[BKPT05] BRESLIN, Thomas ; KROGH, Morten ; PETERSON, Carsten ; TROEIN, Carl:
          Signal transduction pathway profiling of individual tumor samples. In: *BMC
          Bioinformatics* 6 (2005), 163. `https://bscw.icg.tu-graz.ac.at/bscw/`
          `bscw.cgi/d27741/BMCBioinformatics_2005_Signal%20transduction%`
          `20pathway%20profiling%20of%20individual%20tumor%20samples.pdf`

[BR01]   BECKER, Moritz ; ROJAS, Isabel: A Graph Layout Algorithm for
          Drawing Metabolic Pathways. In: *Bioinformatics* 17 (2001), Nr.
          5,   461–467.   `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/`
          `d27670/OxfordPress_Bioinformatics_2001_A%20Graph%20Layout%`
          `20Algorithm%20for%20Drawing%20Metabolic%20Pathways.pdf`

[BSB03]   BEDERSON, Benjamin ; SHNEIDERMAN, Ben ; BEDERSON, Benjamin B.:

The Craft of Information Visualization: Readings and Reflections. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc., 2003. – ISBN 1558609156

[BWK00] BALDONADO, Michelle Q. W. ; WOODRUFF, Allison ; KUCHINSKY, Allan: Guidelines for using multiple views in information visualization. In: *AVI '00: Proceedings of the working conference on Advanced visual interfaces.* ACM Press. – ISBN 1–58113–252–2, 110–119

[CKP+04] CHUNG, Hee-Joon ; KIM, Mingoo ; PARK, Chan H. ; KIM, Jihoon ; KIM, Ju H.: ArrayXPath: mapping and visualizing microarray gene-expression data with integrated biological pathway resources using Scalable Vector Graphics. In: *Nucleic Acids Res* 32 (2004), Jul, Nr. Web Server issue, W460–W464. `http://dx.doi.org/10.1093/nar/gkh476`. – DOI 10.1093/nar/gkh476

[CLRS01] CORMEN, Thomas H. ; LEISERSON, Charles E. ; RIVEST, Ronald L. ; STEIN, Clifford: *Introduction To Algorithms, Second Edition.* Cambridge, MA, USA : MIT Press, 2001. – ISBN 0–262–03293–7

[CPH+05] CHUNG, Hee-Joon ; PARK, Chan H. ; HAN, Mi R. ; LEE, Seokho ; OHN, Jung H. ; KIM, Jihoon ; KIM, Jihun ; KIM, Ju H.: ArrayXPath II: mapping and visualizing microarray gene-expression data with biomedical ontologies and integrated biological pathway resources using Scalable Vector Graphics. In: *Nucleic Acids Res* 33 (2005), Jul, Nr. Web Server issue, W621–W626. `http://dx.doi.org/10.1093/nar/gki450`. – DOI 10.1093/nar/gki450

[DBD+02] DEMIR, E. ; BABUR, O. ; DOGRUSOZ, U. ; GURSOY, A. ; NISANCI, G. ; CETIN-ATALAY, R. ; OZTURK, M.: PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. In: *Bioinformatics* 18 (2002), Nr. 7, 996–1003. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27638/OxfordPress_Bioinformatics_2002_PATIKA%3a%20An%20Integrated%20Visual%20Environment%20For%20Collaborative%20Construction%20and%20Analysis%20Of%20Cellular%20Pathways.pdf`

[DGC+04] DOGRUSOZ, Ugur ; GIRAL, Erhan ; CETINTAS, Ahmet ; CIVRIL, Ali ; DEMIR, Emek: A Compound Graph Layout Algorithm for Biological Pathways. In: PACH, János (Editor): *Graph Drawing, New York, 2004*, Springer, 2004, S. pp. 442–447

[DJ03] DICKERSON JA, Blom K Reinot A Lie J Cruz-Neira C Wurtele ES. . ”.” September. In p.: Using Virtual Reality to Understand Complex Metabolic Networks. In: *Atlantic Symposium on Computational Biology and Genomic Information Systems and Technology, In Press*

[GHJV95] Gamma, Erich ; Helm, Richard ; Johnson, Ralph ; Vlissides, John: *Design patterns: elements of reusable object-oriented software.* Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc., 1995. – ISBN 0–201–63361–2

[GTHC02] Grosu, Paul ; Townsend, Jeffrey P. ; Hartl, Daniel L. ; Cavalieri, Duccio: Pathway Processor: a tool for integrating whole-genome expression results into metabolic networks. In: *Genome Res* 12 (2002), Jul, Nr. 7, 1121-6. `http://dx.doi.org/10.1101/gr.226602`. – DOI 10.1101/gr.226602

[HDLT05] Hibbs, M. A. ; Dirksen, N. C. ; Li, K. ; Troyanskaya, O. G.: Visualization methods for statistical analysis of microarray clusters. In: *BMC Bioinformatics* 6 (2005). `http://dx.doi.org/10.1186/1471-2105-6-115`. – DOI 10.1186/1471–2105–6–115. – ISSN 1471–2105

[HK04] Hauser, Helwig ; Kosara, Robert: Interactive Analysis of High-Dimensional Data using Visualization. In: *Workshop on Robustness for High-dimensional Data (RobHD 2004).* Vorau, Austria, May 2004

[HWH04] Harris, Bobby ; Warner, Rob ; Harris, Robert: *The Definitive Guide to Swt and Jface.* APress, 2004. – ISBN 1590593251

[JM03] Jourdan, Fabien ; Melancon, Guy: Tool for metabolic and regulatory pathways visual analysis, 46-55

[KAJR02] Kurhekar, Manish P. ; Adak, Sudeshna ; Jhunjhunwala, S. ; Raghupathy, K.: Genome-Wide Pathway Analysis and Visualization Using Gene Expression Data. In: *Pacific Symposium on Biocomputing*, 2002, S. 462–473

[KG00] Kanehisa, M. ; Goto, S.: KEGG: kyoto encyclopedia of genes and genomes. In: *Nucleic Acids Research* 28 (2000), January, Nr. 1, 27–30. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27645/OxfordPress_NucleicAcidsResearch_1999_KEGG%3a%20Kyoto%20Encyclopedia%20of%20Genes%20and%20Genomes.pdf`. – ISSN 0305–1048

[KGH+06] Kanehisa, M. ; Goto, S. ; Hattori, M. ; Aoki-Kinoshita, K. F. ; Itoh, M. ; Kawashima, S. ; Katayama, T. ; Araki, M. ; Hirakawa, M.: From genomics to chemical genomics: new developments in KEGG. In: *Nucleic Acids Res* 34 (2006), January, Nr. Database issue. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27778/OxfordPress_NucleicAcidsResearch_2006_From%20genomics%20to%20chemical%20genomics%3a%20new%20developments%20in%20KEGG.pdf`. – ISSN 1362–4962

[KHG03] Kosara, Robert ; Hauser, Helwig ; Gresh, Donna L.: An Interaction

View on Information Visualization. In: *State-of-the-Art Proceedings of EU-ROGRAPHICS 2003 (EG 2003)*, 2003, S. 123–137

[Kin04]   KINCAID, Robert: VistaClara: an interactive visualization for exploratory analysis of DNA microarrays. In: *SAC '04: Proceedings of the 2004 ACM symposium on Applied computing.* New York, NY, USA : ACM Press, 2004. – ISBN 1–58113–812–1, S. 167–174

[KP94a]   KARP, P. ; PALEY, S. M.: Automated Drawing of Metabolic Pathways. In: LIM, H. (Editor) ; CANTOR, C. (Editor) ; ROBBINS, R. (Editor): *Third International Conference on Bioinformatics and Genome Research*

[KP94b]   KARP, Peter D. ; PALEY, Suzanne M.: Representations of Metabolic Knowledge: Pathways. In: *ISMB*, 203-211

[KR93]   KARP, Peter D. ; RILEY, Monica: Representations of Metabolic Knowledge. In: *Proceedings of the 1st International Conference on Intelligent Systems for Molecular Biology*, AAAI Press. – ISBN 0–929280–47–4, 207–215

[KSH04]   KOSARA, R. ; SAHLING, G. ; HAUSER, H.: Linking scientific and information visualization with interactive 3D scatterplots. In: *12th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2004, S. 133–140

[Lin02]   LINDROOS, S.G.E.: Visualizing metabolic pathways: comparative genomics and expression analysis. In: *Proceedings of the IEEE* Bd. 90, 1793 – 1802

[MSH+05] MLECNIK, B. ; SCHEIDELER, M. ; HACKL, H. ; HARTLER, J. ; SANCHEZ-CABO, F. ; TRAJANOSKI, Z.: PathwayExplorer: web service for visualizing high-throughput expression data on biological pathways. In: *Nucleic Acids Research* 33 (2005), July, Nr. Web Server issue. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27785/OxfordPress_NucleicAcidsResearch_2005_PathwayExplorer%3a%20web%20service%20for%20visualizing%20high-throughput%20expression%20data%20on%20biological%20pathways.pdf`. – ISSN 1362–4962

[MW95]   MARTIN, Allen R. ; WARD, Matthew O.: High Dimensional Brushing for Interactive Exploration of Multivariate Data. In: *VIS '95: Proceedings of the 6th conference on Visualization '95.* Washington, DC, USA : IEEE Computer Society, 1995. – ISBN 0–8186–7187–4, S. 271

[NBK+99] NAKAO, Mitsuteru ; BONO, Hidemasa ; KAWASHIMA, Shuichi ; KAMIYA, Tomomi ; SATO, Kazushige ; GOTO, Susumu ; KANEHISA, Minoru: Genome-scale Gene Expression Analysis and Pathway Reconstruction in KEGG. In: *Genome Informatics* Bd. 10, 1999, S. 94–103

[NK02] NIEMEYER, Pat ; KNUDSEN, Jonathan: *Learning Java.* Sebastopol, CA, USA : O'Reilly & Associates, Inc., 2002. – ISBN 0596002858

[PC06] PELLEGRINI, M. ; COKUS, S.: Pathway Analysis of Microarray Data. In: *In Press* (2006)

[RC02] ROJDESTVENSKI, Igor ; COTTAM, Michael: Visualizing Metabolic Networks in VRML. In: *6th International Conference on InformationVisualisation* 00 (2002), 175. `http://dx.doi.org/http://doi.ieeecomputersociety.org/10.1109/IV.2002.1028774`. – DOI http://doi.ieeecomputersociety.org/10.1109/IV.2002.1028774. – ISSN 1093–9547

[RCB+05] ROBERTSON, George ; CZERWINSKI, Mary ; BAUDISCH, Patrick ; MEYERS, Brian ; ROBBINS, Daniel ; SMITH, Greg ; TAN, Desney: The Large-Display User Experience. In: *IEEE Comput. Graph. Appl.* 25 (2005), Nr. 4, 44–51. `http://dx.doi.org/http://dx.doi.org/10.1109/MCG.2005.88`. – DOI http://dx.doi.org/10.1109/MCG.2005.88. – ISSN 0272–1716

[Roj03] ROJDESTVENSKI, Igor: Metabolic pathways in three dimensions. In: *Bioinformatics* 19 (2003), Nr. 18, 2436-2441. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27662/OxfordPress_Bioinformatics_2003_Metabolic%20Pathways%20In%20Three%20Dimensions.pdf`

[Sch03a] SCHREIBER, Falk: Comparison of metabolic pathways using constraint graph drawing. In: *APBC '03: Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003.* Australian Computer Society, Inc.. – ISBN 0–909–92597–6, 105–110

[Sch03b] SCHREIBER, Falk: Visual Comparison of Metabolic Pathways. In: *Journal of Visual Languages and Computing* 14 (2003), Nr. 4, 327–340. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27761/2003_Visual%20Comparison%20Of%20Metabolic%20Pathways.pdf`

[Shn96] SHNEIDERMAN, Benn: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: *VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages*, IEEE Computer Society. – ISBN 081867508X

[SM00] SCHUMANN, Heidrun ; MÃ¼LLER, Wolfgang: *Visualisierung - Grundlagen und allgemeine Methoden.* Berlin Heidelberg, Germany : Springer Verlag, 2000. – 370 S

[Spe07] SPENCE, Robert: *Information Visualization: Design for Interaction (2nd Edition).* Prentice-Hall, Inc., 2007. – ISBN 978–0–132–0655–04

[SS02]   SEO, Jinwook ; SHNEIDERMAN, Ben: Interactively Exploring Hierarchical Clustering Results. In: *Computer* 35 (2002), Nr. 7, 80–86. `http://dx.doi.org/http://dx.doi.org/10.1109/MC.2002.1016905`. –   DOI http://dx.doi.org/10.1109/MC.2002.1016905. – ISSN 0018–9162

[SWND05]   SHREINER, Dave ; WOO, Mason ; NEIDER, Jackie ; DAVIS, Tom: *OpenGL Programming Guide Fifth Edition: The Official Guide to Learning OpenGL, Version 2.* 5. Addison-Wesley Professional, 2005. – 896 S. – ISBN 978–0321335739

[TSB04]   TANG, Diane ; STOLTE, Chris ; BOSCH, Robert: Design choices when architecting visualizations. In: *Information Visualization* 3 (2004), Nr. 2, 65–79. `http://dx.doi.org/http://dx.doi.org/10.1057/palgrave.ivs.9500067`. –   DOI http://dx.doi.org/10.1057/palgrave.ivs.9500067. – ISSN 1473–8716

[TSKS05]   TANAY, Amos ; STEINFELD, Israel ; KUPIEC, Martin ; SHAMIR, Ron: Integrative analysis of genome-wide experiments in the context of a large high-throughput data compendium. In: *Molecular Systems Biology* 1 (2005), March, Nr. 1, S. msb4100005–E1–msb4100005–E10. `http://dx.doi.org/10.1038/msb4100005`. – DOI 10.1038/msb4100005

[WF94]   WARE, Colin ; FRANCK, Glenn: Viewing a Graph in a Virtual Reality Display is Three Times as Good as 2D Diagram. In: *VL*, 1994, S. 182–183

[WGS00]   WOLF, Detlef ; GRAY, Christopher P. ; SAIZIEU, Antoine de: Visualising Gene Expression in Its Metabolic Context. In: *Briefings in Bioinformatics* 1 (2000), Nr. 3, 297-304. `https://bscw.icg.tu-graz.ac.at/bscw/bscw.cgi/d27725/2000_Visualising%20Gene%20Expression%20In%20Its%20Metabolic%20Context.pdf`

[Wol05]   WOLFF, David: Using openGL in Java with JOGL. In: *J. Comput. Small Coll.* 21 (2005), Nr. 1, S. 223–224

[XYC05]   XU, Zhigen ; YAN, Yusong ; CHEN, Jim X.:   OpenGL Programming in Java. In: *Computing in Science and Engg.* 7 (2005), Nr. 1, S. 51–55. `http://dx.doi.org/http://dx.doi.org/10.1109/MCSE.2005.20`. – DOI http://dx.doi.org/10.1109/MCSE.2005.20. – ISSN 1521–9615

[YEW⁺05]   YANG, Yuting ; ENGIN, Levent ; WURTELE, Eve S. ; CRUZ-NEIRA, Carolina ; DICKERSON, Julie A.:   Integration of metabolic networks and gene expression in virtual reality. In: *Bioinformatics* 21 (2005), Nr. 18, 3645–3650. `http://dx.doi.org/http://dx.doi.org/10.1093/bioinformatics/bti581`. –   DOI http://dx.doi.org/10.1093/bioinformatics/bti581. – ISSN 1367–4803

[YWCND06] YANG, Yuting ; WURTELE, Eve S. ; CRUZ-NEIRA, Carolina ; DICKERSON, Julie A.: Hierarchical visualization of metabolic networks using virtual reality. In: *VRCIA '06: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications.* ACM Press. – ISBN 1–59593–324–7, 377–381

Index

bla, 32