Design Studies

Lecture 3 CPSC 533C, Fall 2005

Mon Sep 19 2005

Tamara Munzner

Questions

5 questions for the day, not 5 per reading!

ASCII text not Word attachments

audience: up to you

- author, professor, yourself, classmates, ...

Papers Covered

Cluster and Calendar based Visualization of Time Series Data.
Jarke J. van Wijk and Edward R. van Selow, pp 4–9
Proc. InfoVis 99.
http://www.win.tue.nl/~vanwijk/clv.pdf

Using Multilevel Call Matrices in Large Software Projects.
Frank van Ham,
Proc. InfoVis 2003, pp 227–232
http://www.win.tue.nl/~fvham/DL/callmatrix.pdf

Constellation: Linguistic Semantic Networks
Tamara Munzner,
Interactive Visualization of Large Graphs and Networks (PhD thesis) Chapter 5, Stanford University, 2000, pp 87–122
http://graphics.stanford.edu/papers/munzner_thesis

Design Study

describe task

justify solution

refine until satisfied

Design Study Definition

Design study papers explore the choices made when applying infovis techniques in an application area, for example relating the visual encodings and interaction techniques to the requirements of the target task. Although a limited amount of application domain background information can be useful to provide a framing context in which to discuss the specifics of the target task, the primary focus of the case study must be the infovis content. Describing new techniques and algorithms developed to solve the target problem will strengthen a design study paper, but the requirements for novelty are less stringent than in a Technique paper.

InfoVis03 CFP, [infovis.org/infovis2003/CFP]

Time-series Data Analysis

data: N pairs of (value, time)

- N large: 50K

tasks

- find standard day patterns
- find how patterns distributed over year, week, season
- find outliers from standard daily patterns
- want overview first, then detail on demand

possibilities

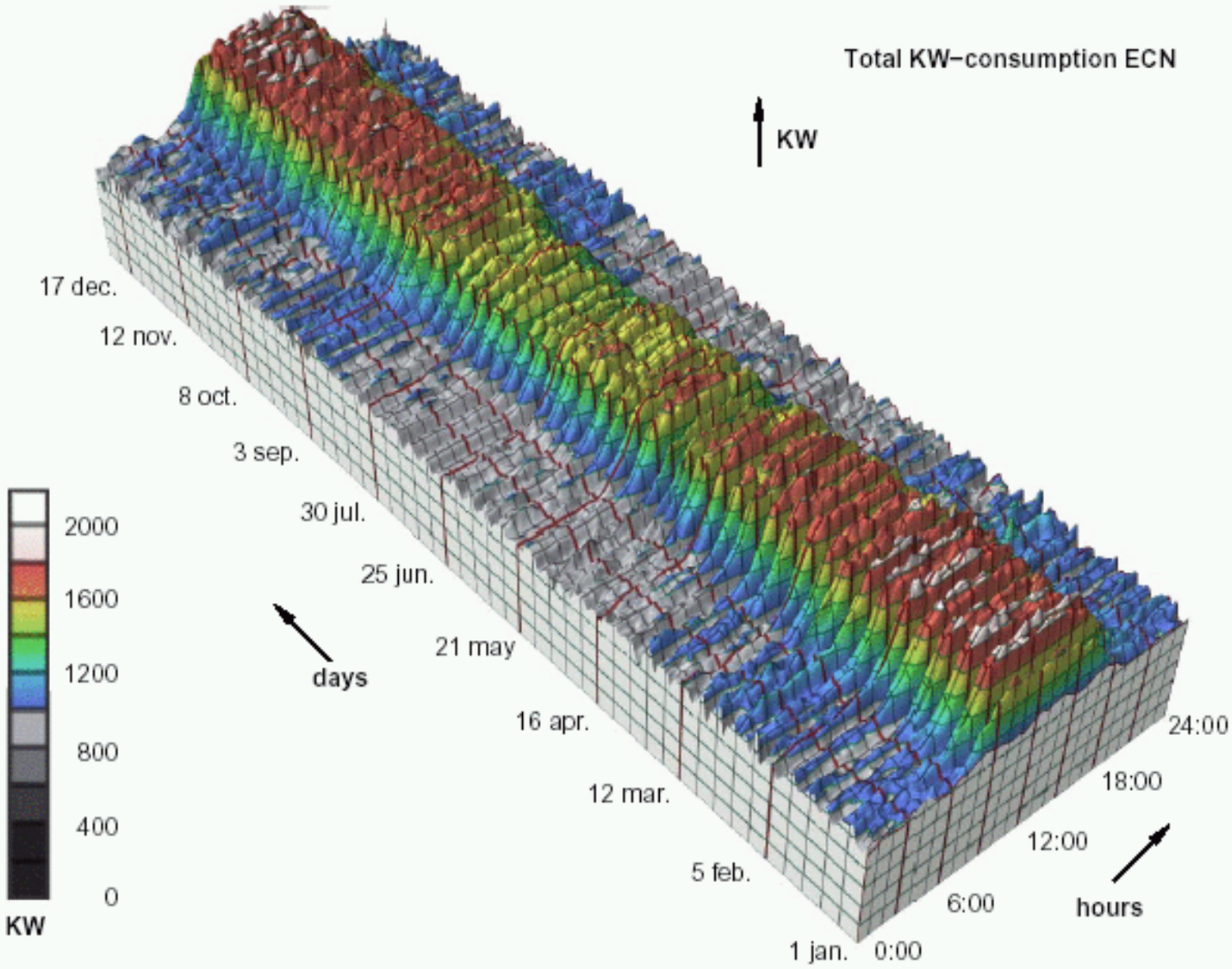
- predictive mathematical models
  details lost, multiscale not addressed
- scale-space approaches (wavelet, fourier, fractal)
  hard to interpret, known scales lost
- 3D mountain: x hours, y value, z days

excellent example, emulate for project writeups!

3D Time-series Data

3D extrusion pretty but not useful

- daily, weekly patterns hard to see



[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, http://www.win.tue.nl/~vanwijk/clv.pdf]

Hierarchical Clustering

start with all M day patterns

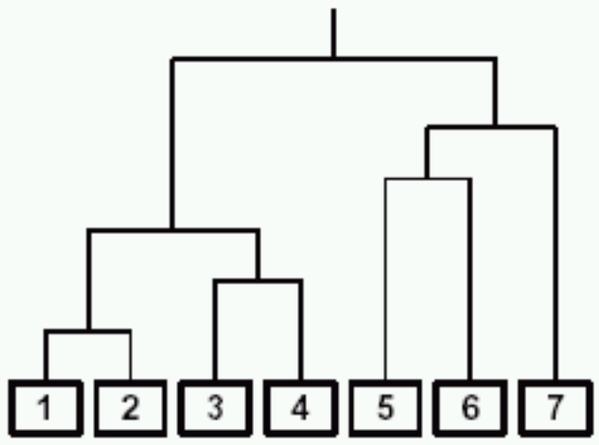
- compute mutual differences, merge most similar: M-1
- continue up to 1 root cluster

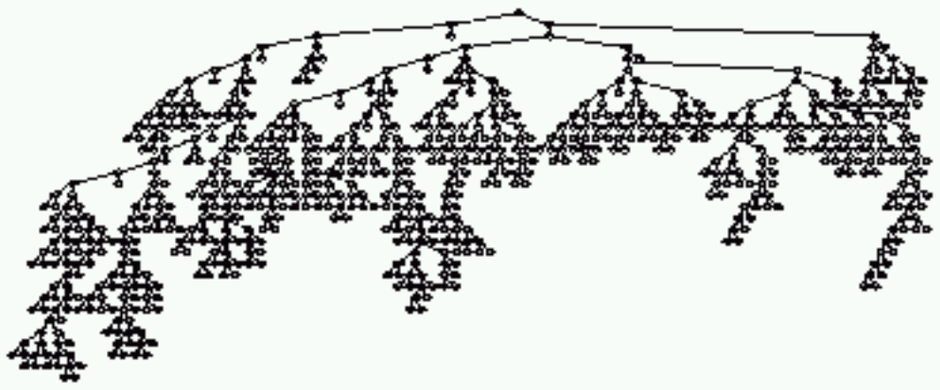
result: binary hierarchy of clusters

choice of distance metrics

dendrogram display common

- but shows structure of hierarchy, not time distribution

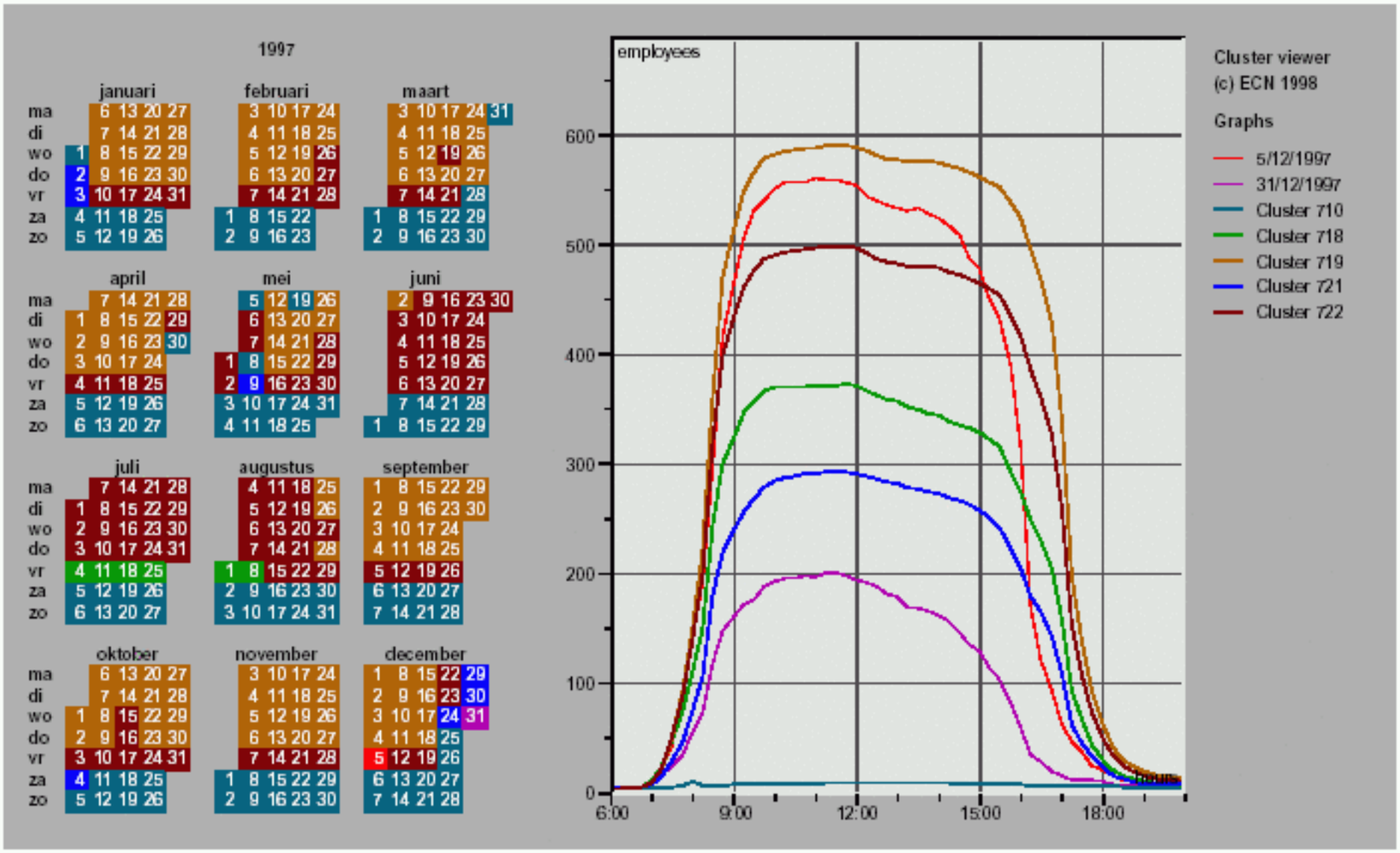




Link Clusters and Calendar

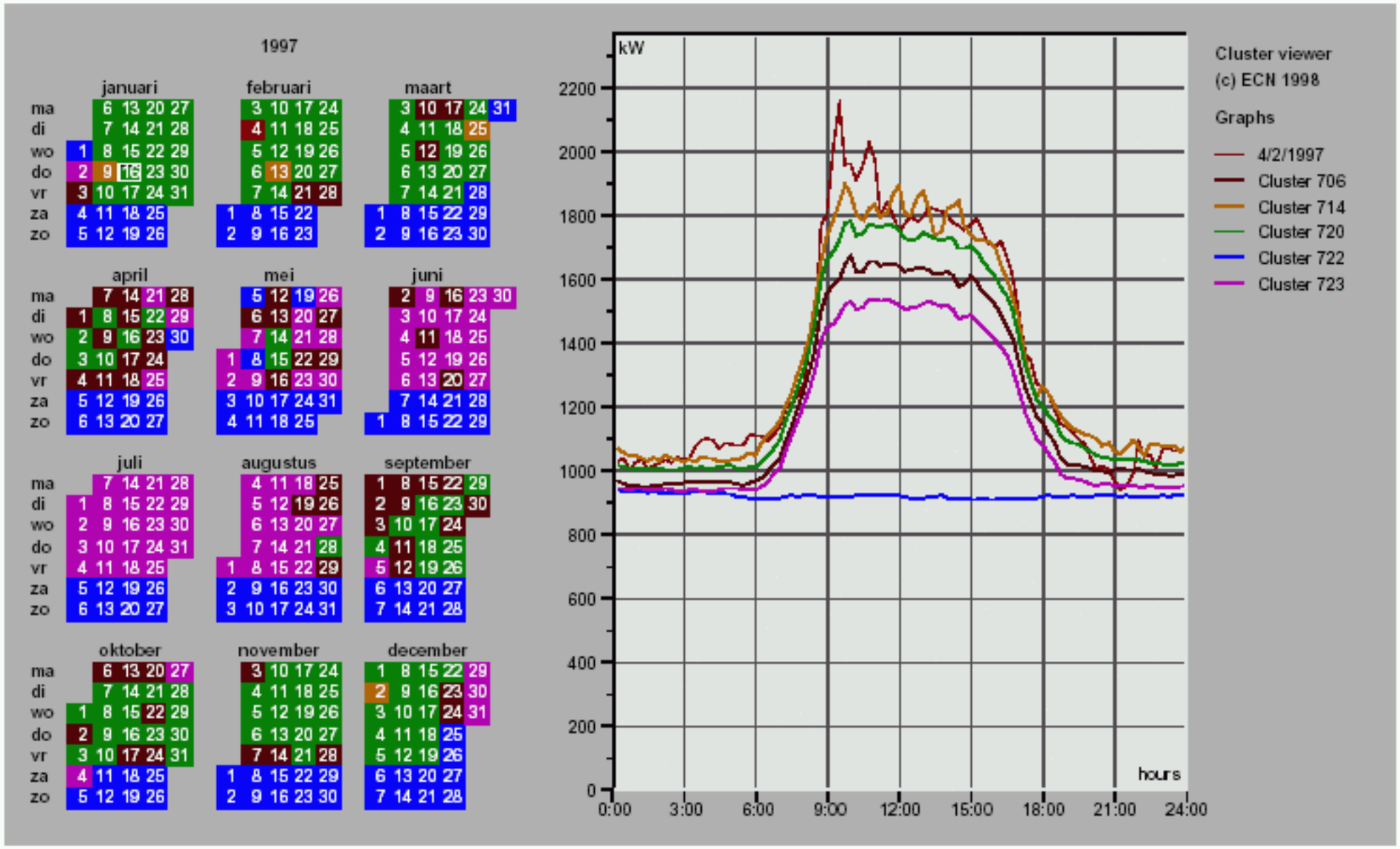
2D linked clusters–calendars shows patterns

- number of employees:
- office hours, fridays in/and summer, school break
- weekend/holidays, post-holiday, santa claus



[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, Figure 4, http://www.win.tue.nl/~vanwijk/clv.pdf]

Power Consumption



[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, Figure 5, http://www.win.tue.nl/~vanwijk/clv.pdf]

van Wijk Lessons

derived space: clusters
visual representation of time: calendar
- linked display
- interactive exploration

clear task analysis guided choices
- reject standard 3D extrusion
- reject standard dendrogram

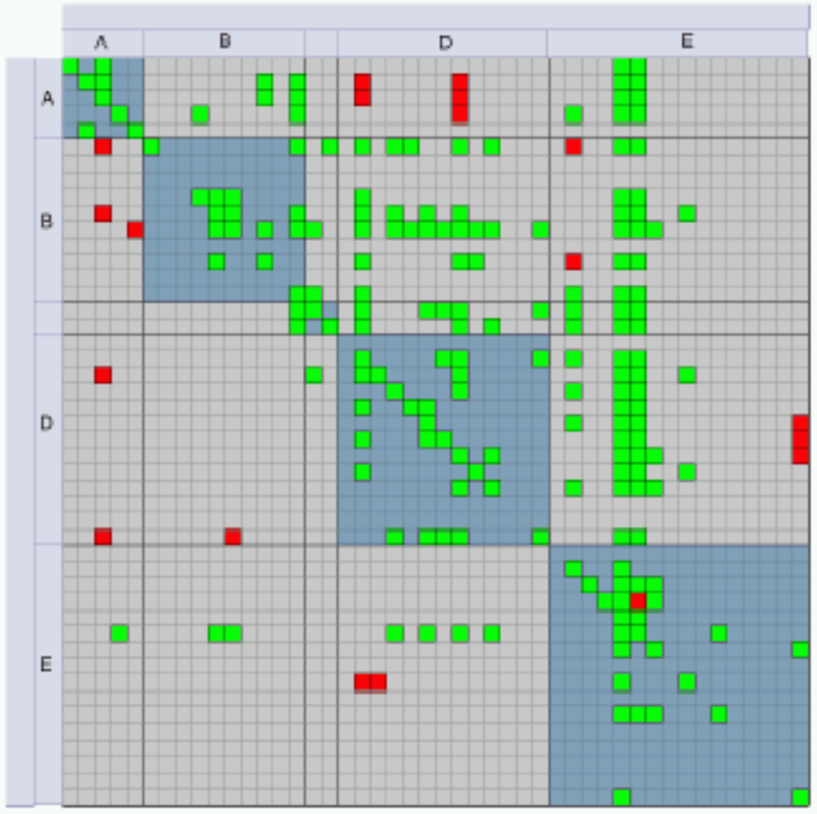
critique
- color choice not so discriminable
  especially legend

Multilevel Call Matrices, van Ham

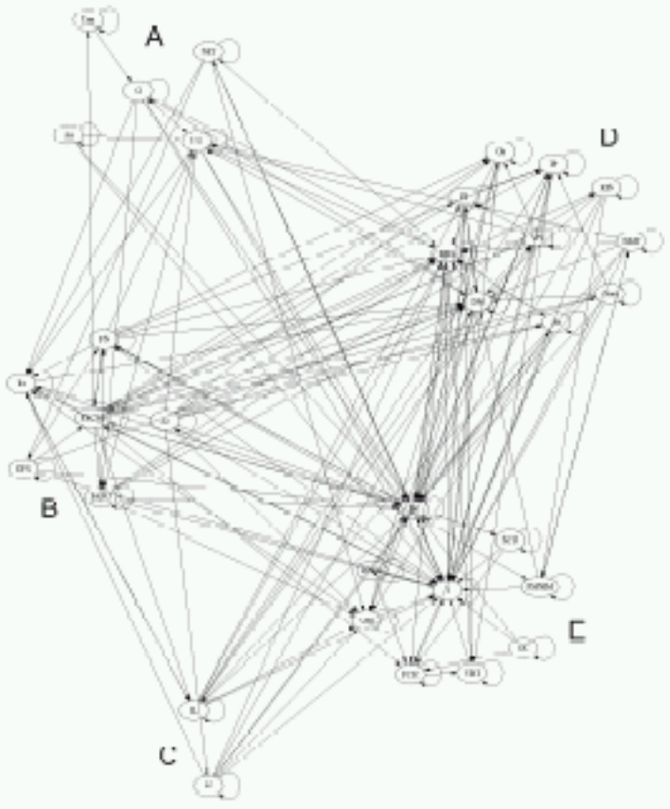
large software project, implementation vs. spec

link matrix vs. node network

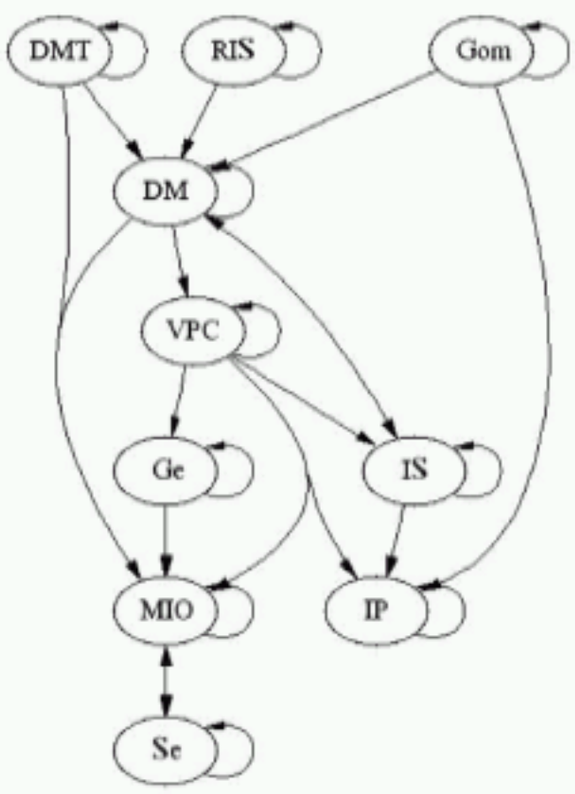
matrix



force-directed



layered subset



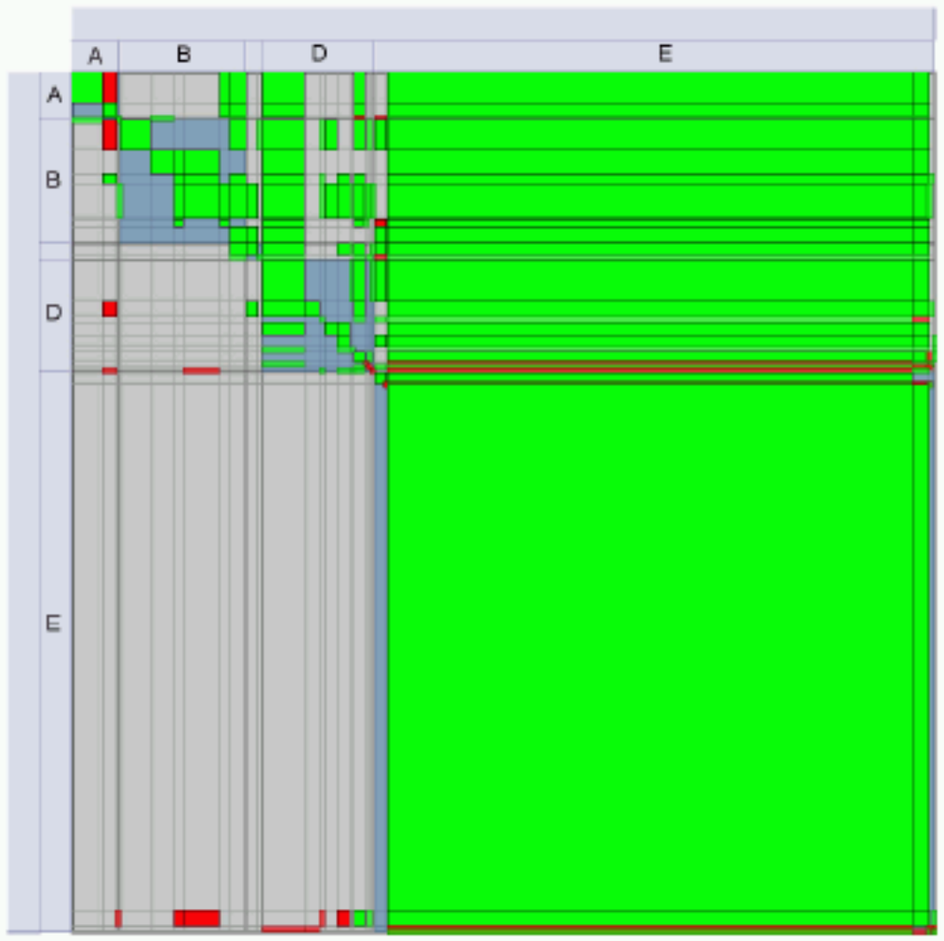
[van Ham, Using Multilevel Call Matrices in Large Software Projects. InfoVis03
http://www.win.tue.nl/~fvham/DL/callmatrix.pdf]

Matrices

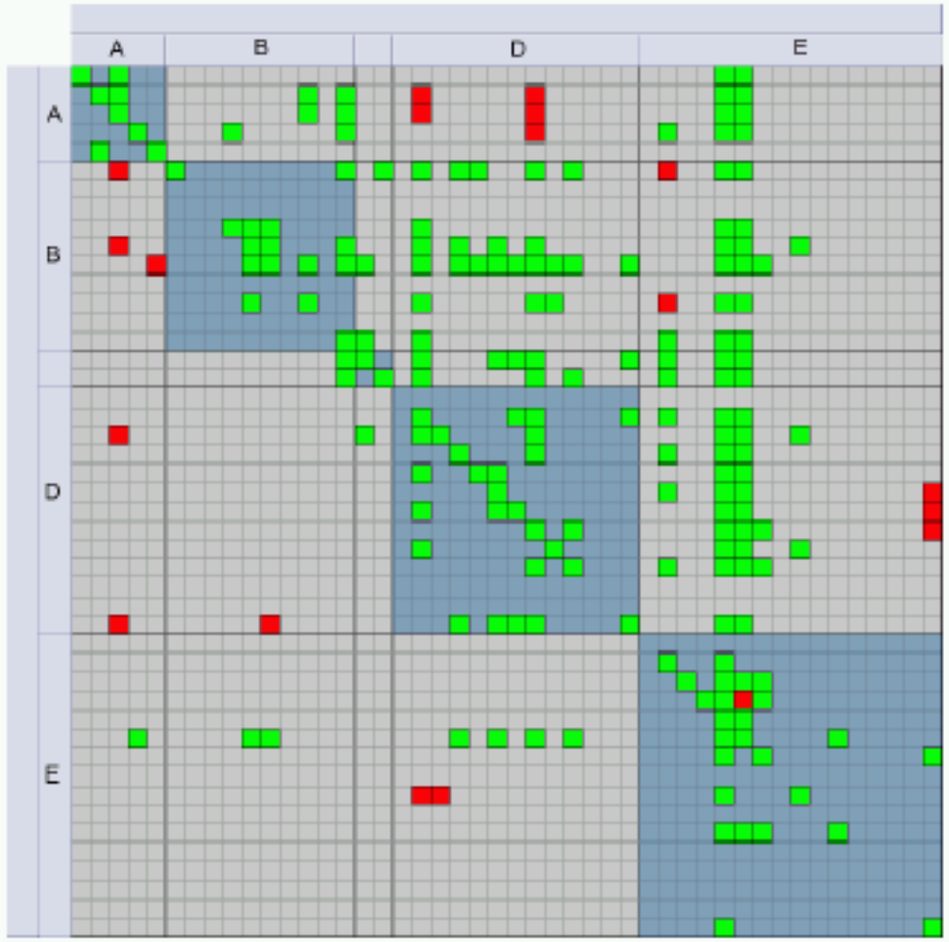
uniform, recursive, stable

subdivide by

total component count



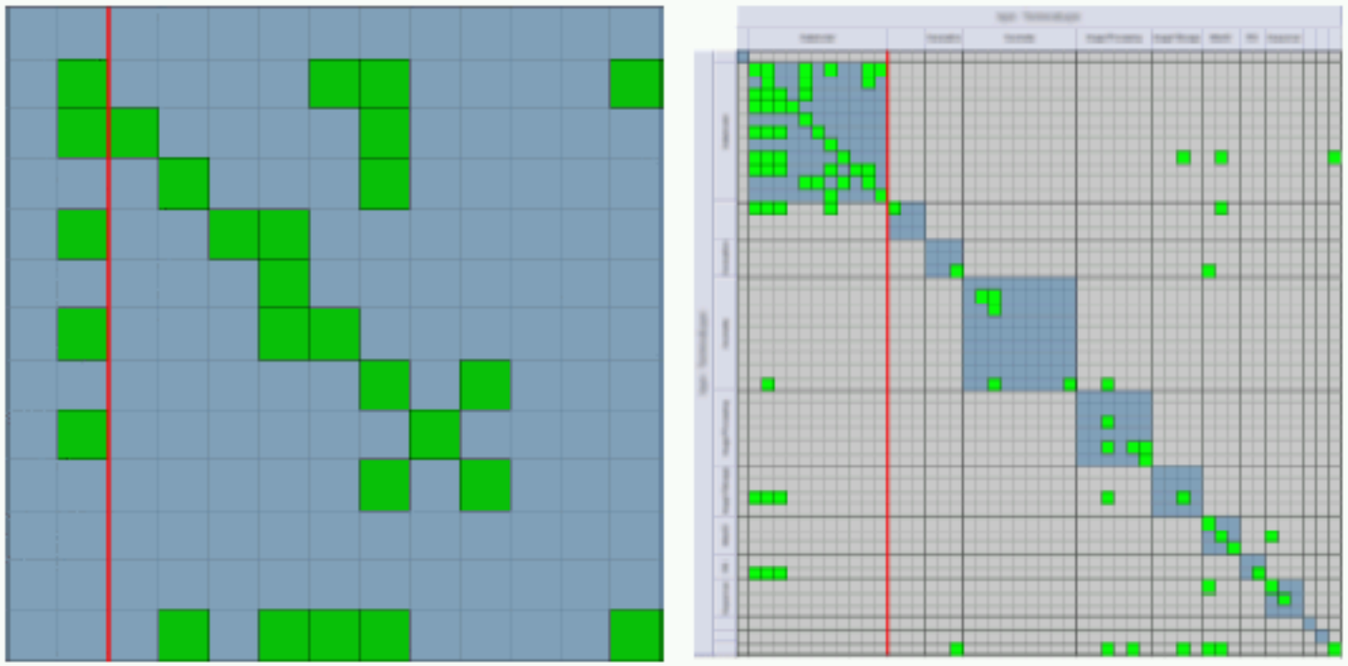
visible subcomponent count



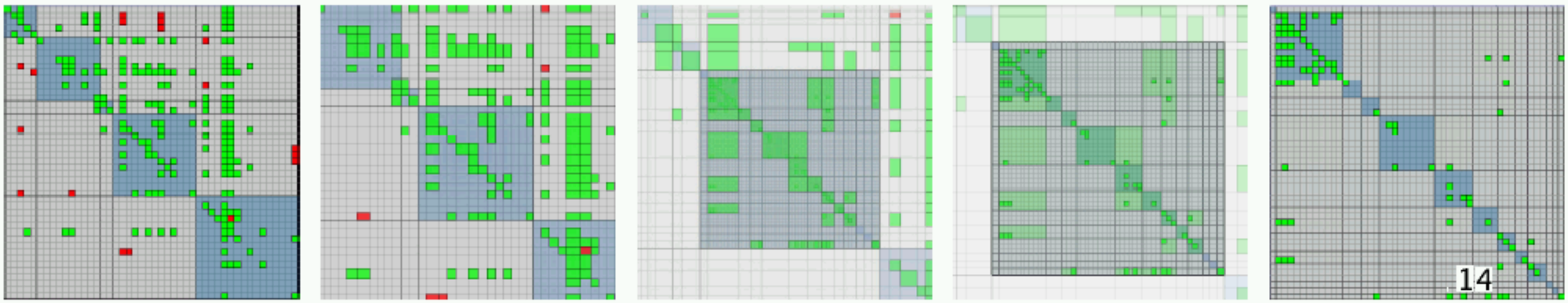
[van Ham, Using Multilevel Call Matrices in Large Software Projects. InfoVis03
http://www.win.tue.nl/~fvham/DL/callmatrix.pdf]

Zooming

abstraction levels



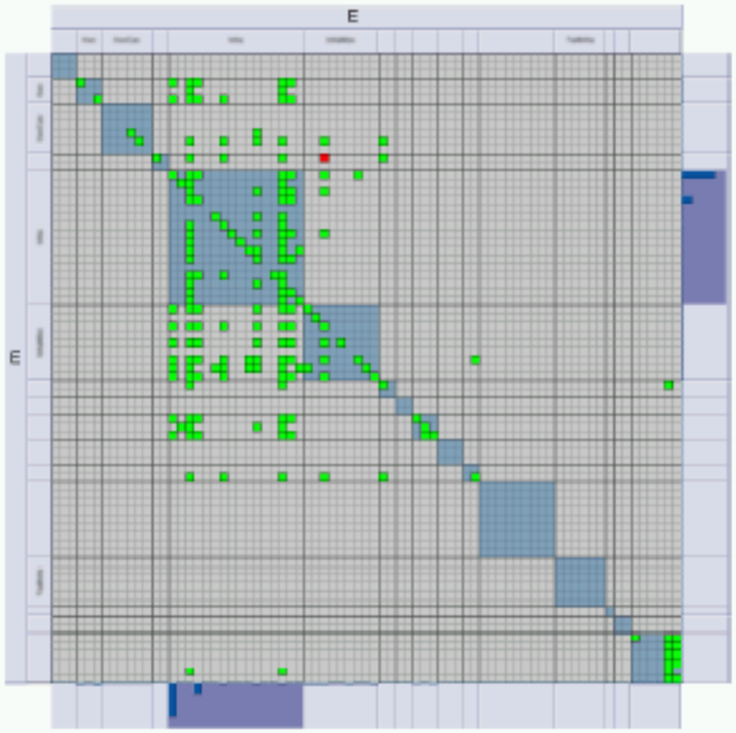
linear interpolation plus crossfade
trajectories: will read van Wijk 03 in week 6



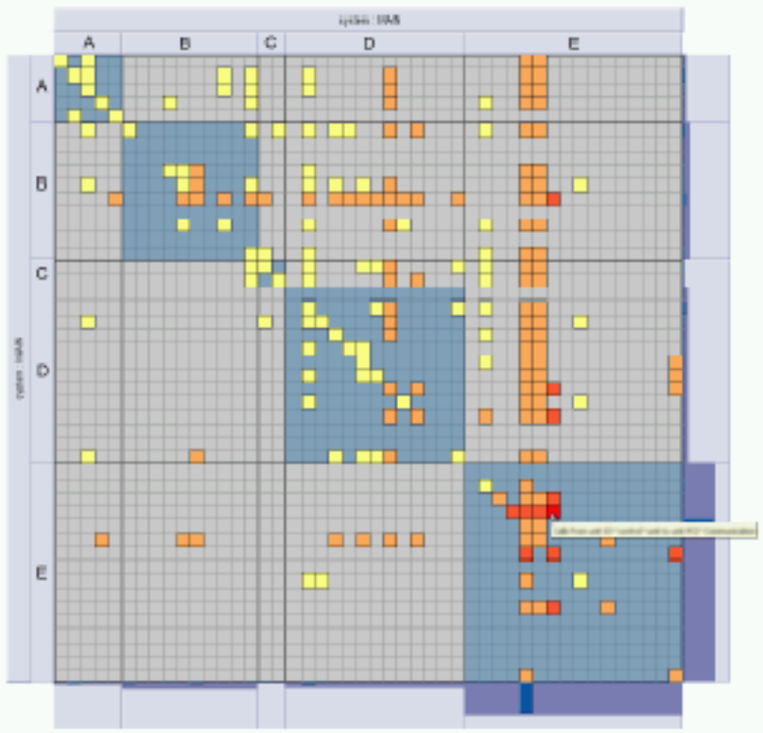
[van Ham, Using Multilevel Call Matrices in Large Software Projects. InfoVis03
http://www.win.tue.nl/~fvham/DL/callmatrix.pdf]

Additional Encoding

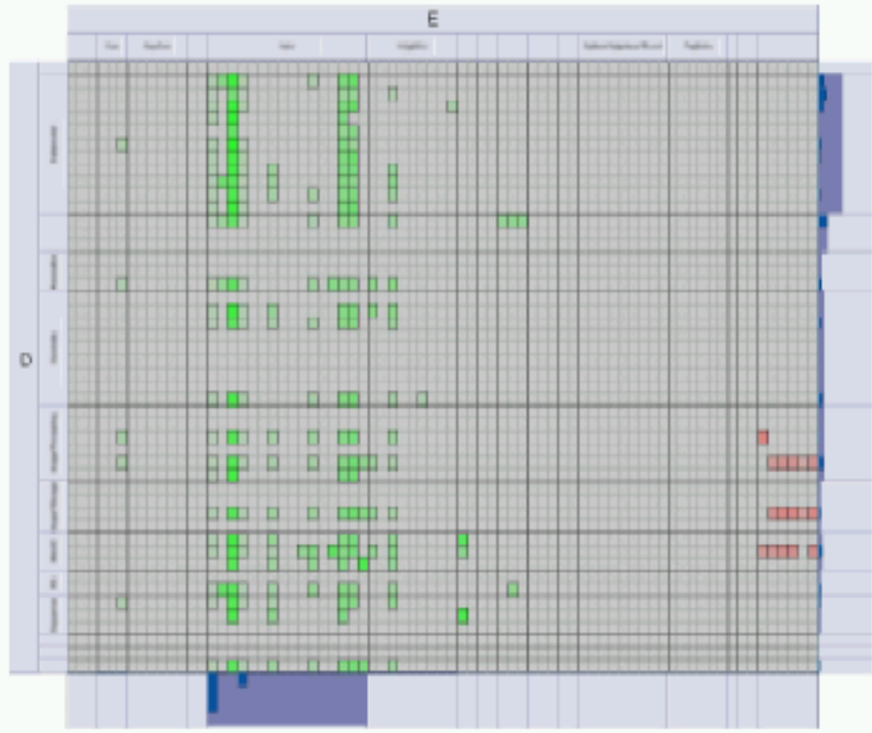
color:
call allowed
by spec



color:
local region
closest red



transparency:
call density



histograms: size distribution

[van Ham, Using Multilevel Call Matrices in Large Software Projects. InfoVis03
http://www.win.tue.nl/~fvham/DL/callmatrix.pdf]

Tasks Succesfully Supported

visual categorization

- i.e. libraries with mostly incoming calls

previous summary shown to be incomplete

spotting unwanted calls

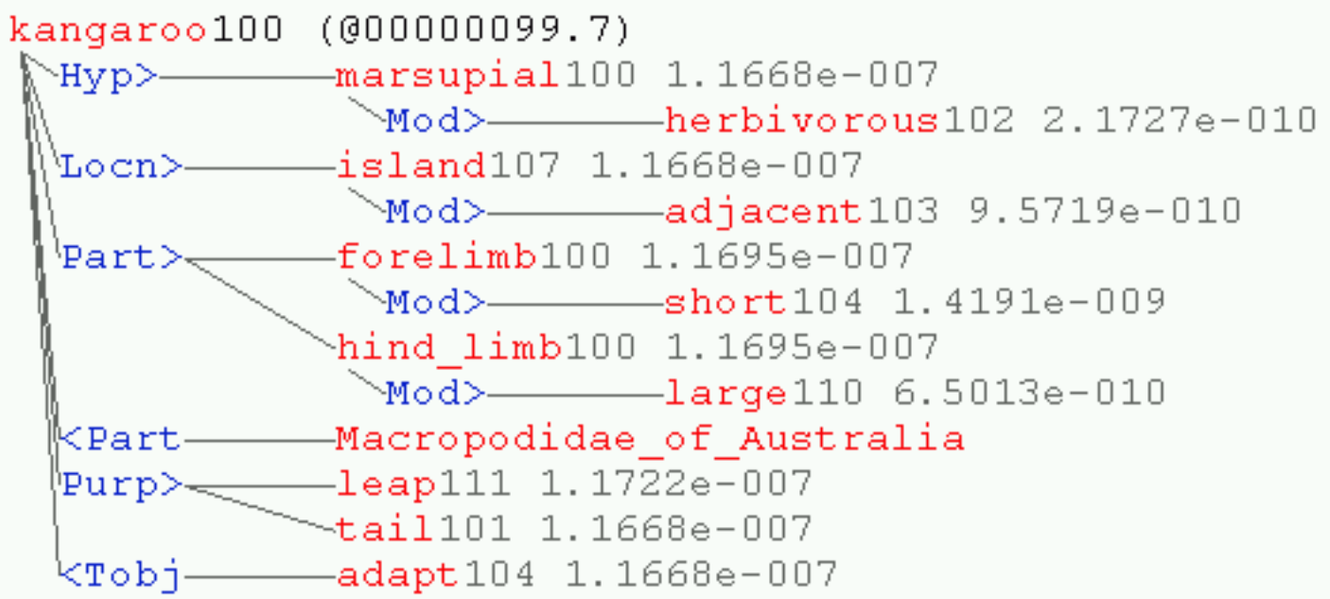
determining component dependencies

Linguistic Networks, Munzner

data: MindNet query results

definition graph

- dictionary entry sentence
- nodes: word senses
- links: relation types



Semantic Network

definition graphs used as building blocks

unify shared words

large network

- millions of nodes
- grammar checking now, translation future
- global structure known: dense

probes return local info

Path Query

best N paths between two words

words on path itself

```
kangaroo100—Part→forelimb100—Mod→short104—Join→short←Mod—tail100
```

definition graphs used in computation

```
kangaroo100  (vole101 tapir100 s
sharp-tailed_grouse100 scut100 r
pitta100 partridge104 lynx100 lo
kingfisher100 horned_toad100 haw
bobtail101 bobtail100 bobcat100
Scottish_terrier100)
```

Task: Plausibility Checking

paths ordered by computed plausibility

researcher hand-checks results

- high-ranking paths believable?
- believable paths high-ranked?
- are stop words all filtered out?

Top 10 Paths Kangaroo->Tail

Natural Language Processor [Ansi, Debug, BigSys]

File Analyze... Command... Display Choose Explain Tools Options Window Help

Graph - Path - "kangaroo" "tail"

```
Number of paths: 10


Similarity score: 0.00068368 ( < 0.0015 - the words are not similar)

1       1.1668e-007  kangaroo100—Purp->tail101  kangaroo100

2       6.4417e-014  kangaroo100—Hyp->marsupial100<-Hyp—Tasmanian_devil100—Part->tail101  kangaroo100  Tasmanian_devil100

3       4.9545e-014  kangaroo103—Hyp->animal109—Part->tail101  kangaroo103  (taper103 tail127 tail111
                                                               tag114 switch115 dock111 chipmunk102)

4       4.2954e-014  kangaroo100—Hyp->marsupial100<-Hyp—cuscus100—Part->tail101  kangaroo100  cuscus100

5       1.2972e-014  kangaroo100—Part->forelimb100—Mod->short104—Join->short<-Mod—tail100  kangaroo100  (vole101 tapir100 s
                                                                                             sharp-tailed_grouse100 scut100 r
                                                                                             pitta100 partridge104 lynx100 lo
                                                                                             kingfisher100 horned_toad100 haw
                                                                                             bobtail101 bobtail100 bobcat100
                                                                                             Scottish_terrier100)

6       5.6234e-015  kangaroo103<-Hyp—wallaroo100—Part->fur112—Join->fur113—Mod->tail132  wallaroo100  (phalanger100 ermine

7       2.4774e-015                                                                        il111
                                                                                           munk102)

8       1.5560e-015                                                                        Old_English_sheepdo

9       1.5488e-015                                                                        wolverine100

10      1.1220e-015                                                                        aby100  rock_wallaby
```

```
kangaroo100 (@00000099.7)
  Hyp>———marsupial100 1.1668e-007
          Mod>———herbivorous102 2.1727e-010
  Locn>———island107 1.1668e-007
          Mod>———adjacent103 9.5719e-010
  Part>———forelimb100 1.1695e-007
          Mod>———short104 1.4191e-009
        hind_limb100 1.1695e-007
          Mod>———large110 6.5013e-010
  <Part———Macropodidae_of_Australia
  Purp>———leap111 1.1722e-007
        tail101 1.1668e-007
  <Tobj———adapt104 1.1668e-007
```

Goal

create a unified view of relationships between paths and definition graphs

- shared words are key
- thousands of words (not millions)

special purpose algorithm debugging tools

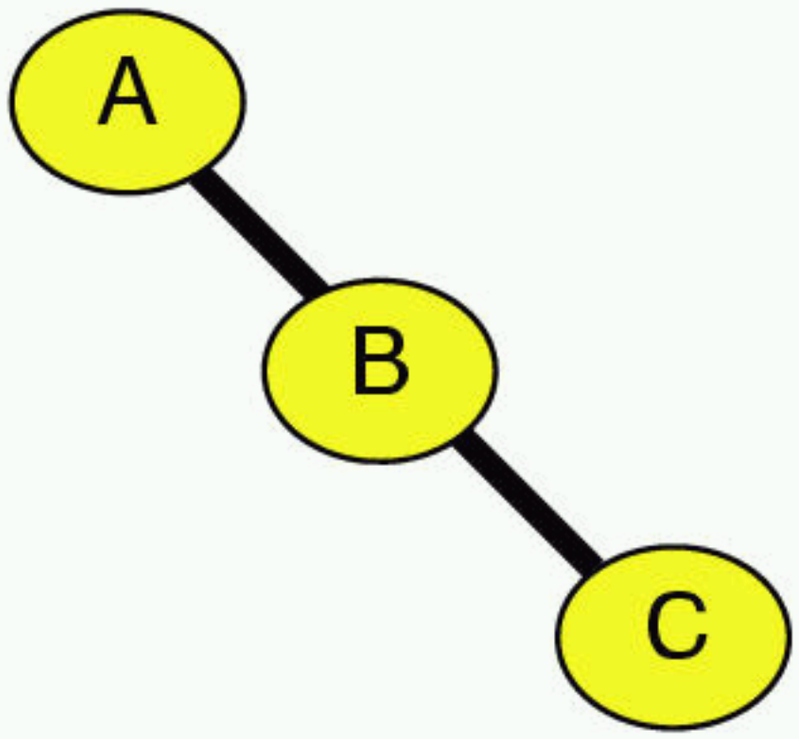
- not understand structure of English

Constellation Video

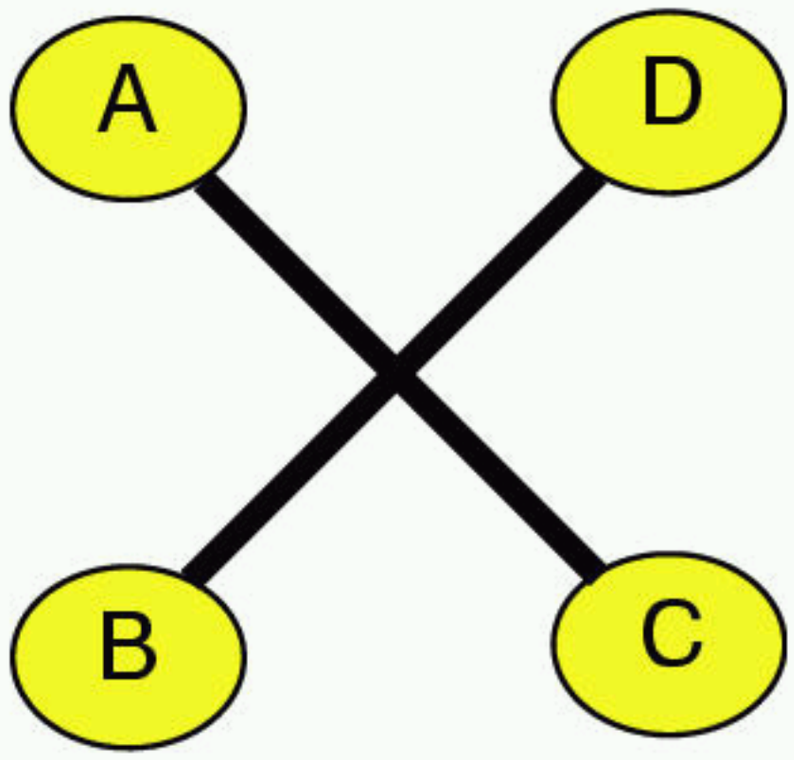
Traditional Layout

avoid crossings

reason: avoid false attachments



ambiguity

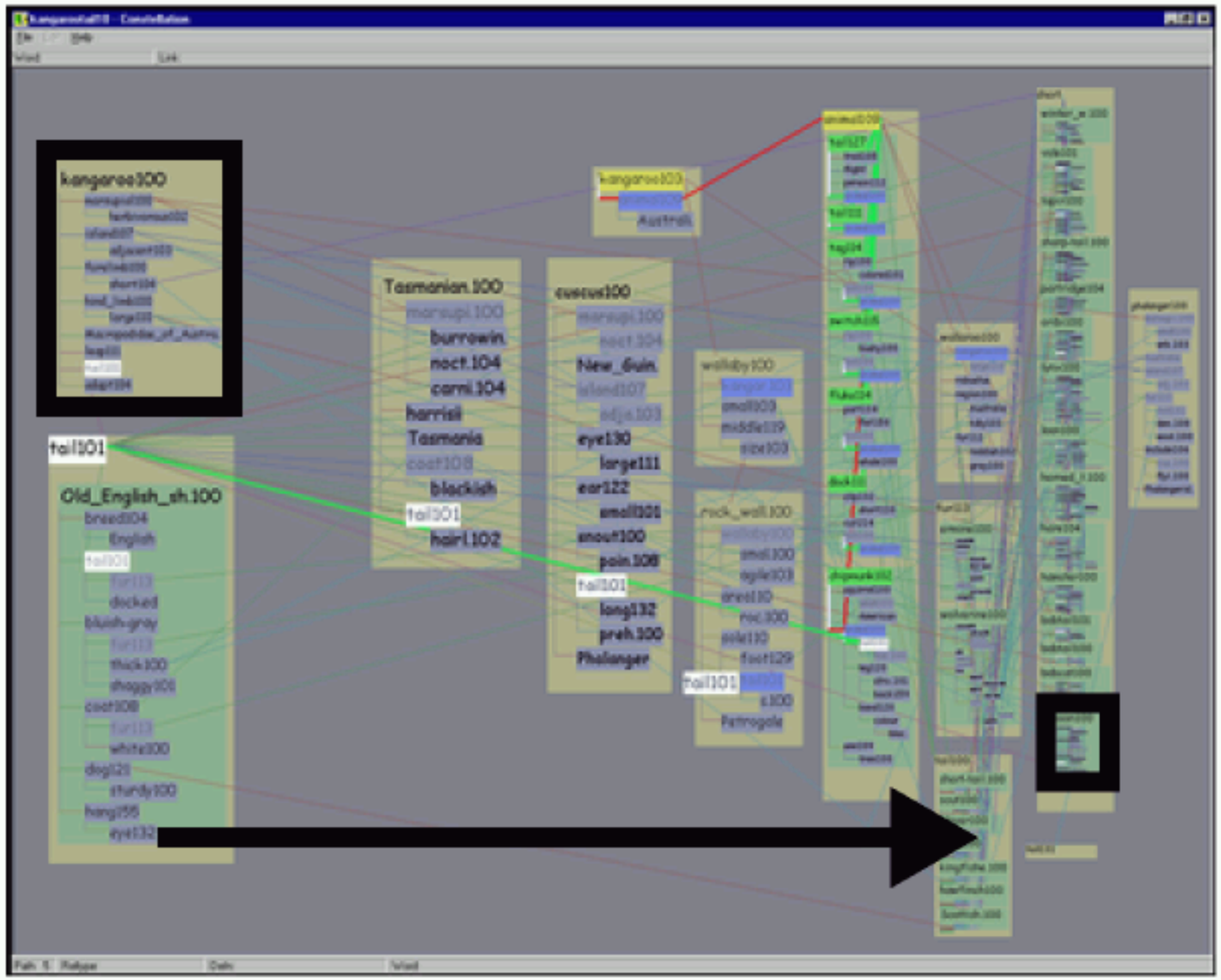


artifact salience

Information Visualization Approach

spatial position is strongest perceptual cue

- encode domain specific attribute
- plausibility gradient



Constellation Semantic Layout

novel layout algorithm

- paths as backbone, definition graphs attached
- curvilinear grid
- iterative design for maximum semantics with reasonable information density

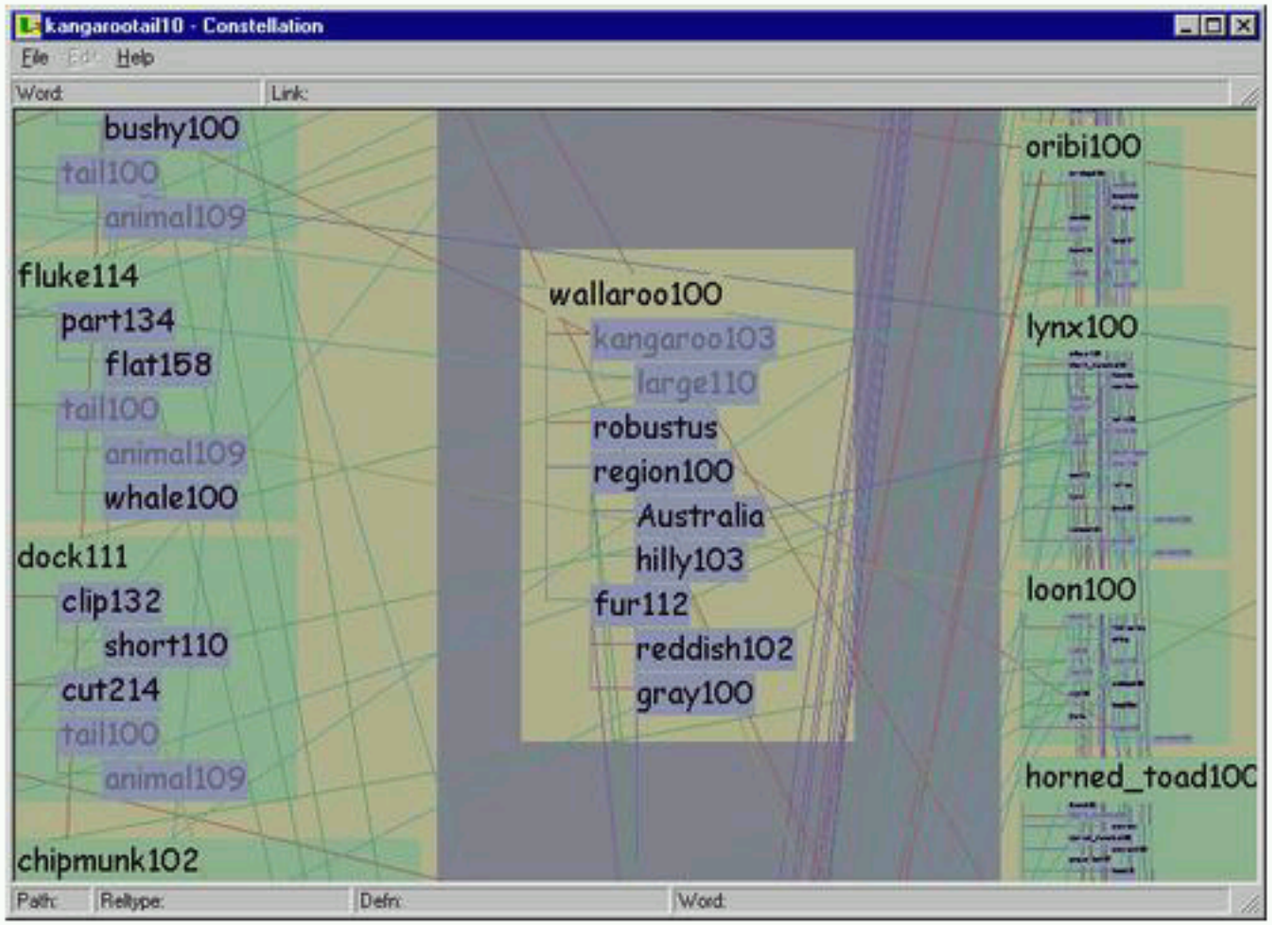
allow crossings for long-distance proxy links

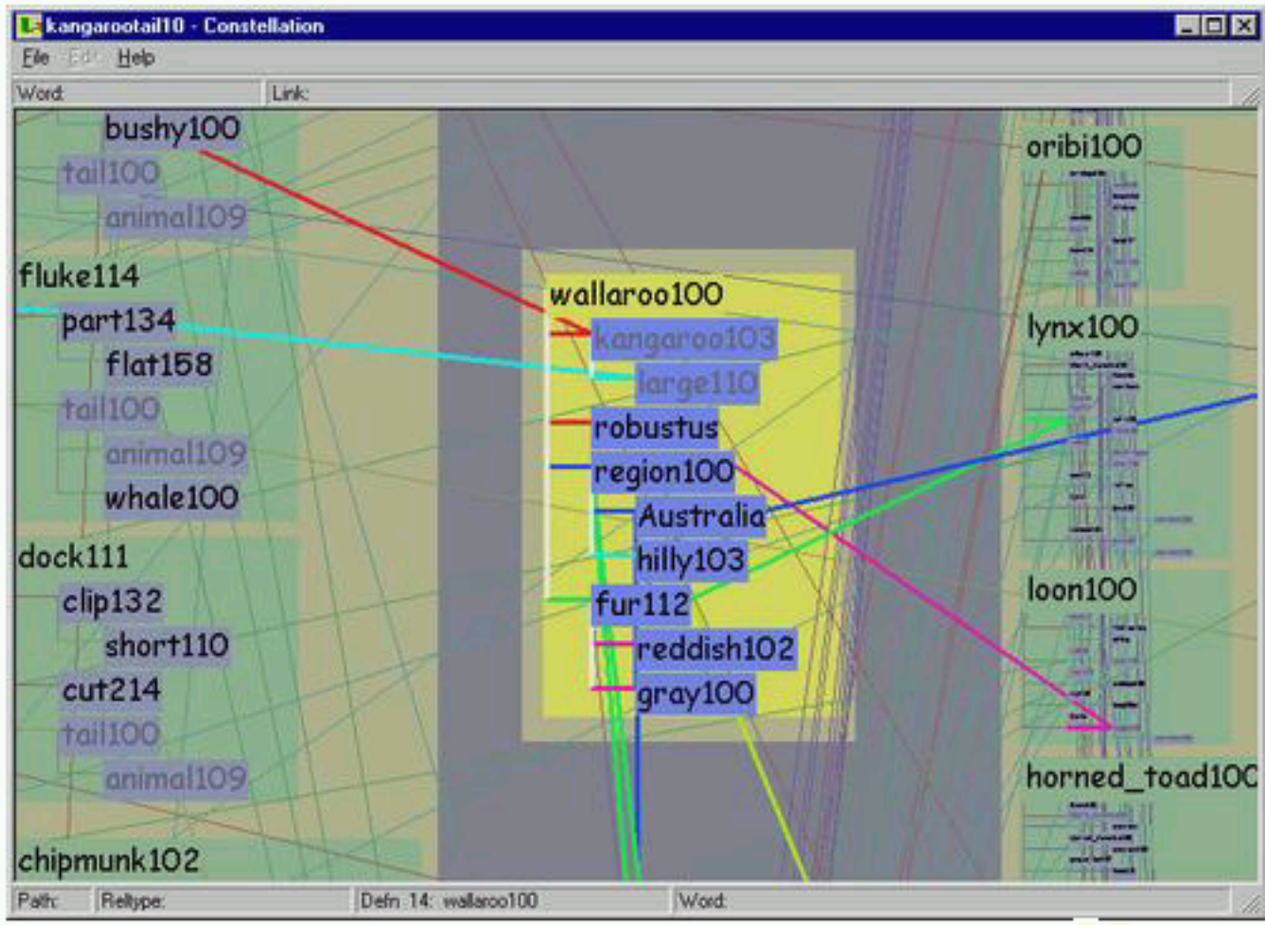
Selective Emphasis

highlight sets of boxes and edges

- interaction
- additional perceptual channels

avoid perception of false attachments





Hidden State

avoid hidden state

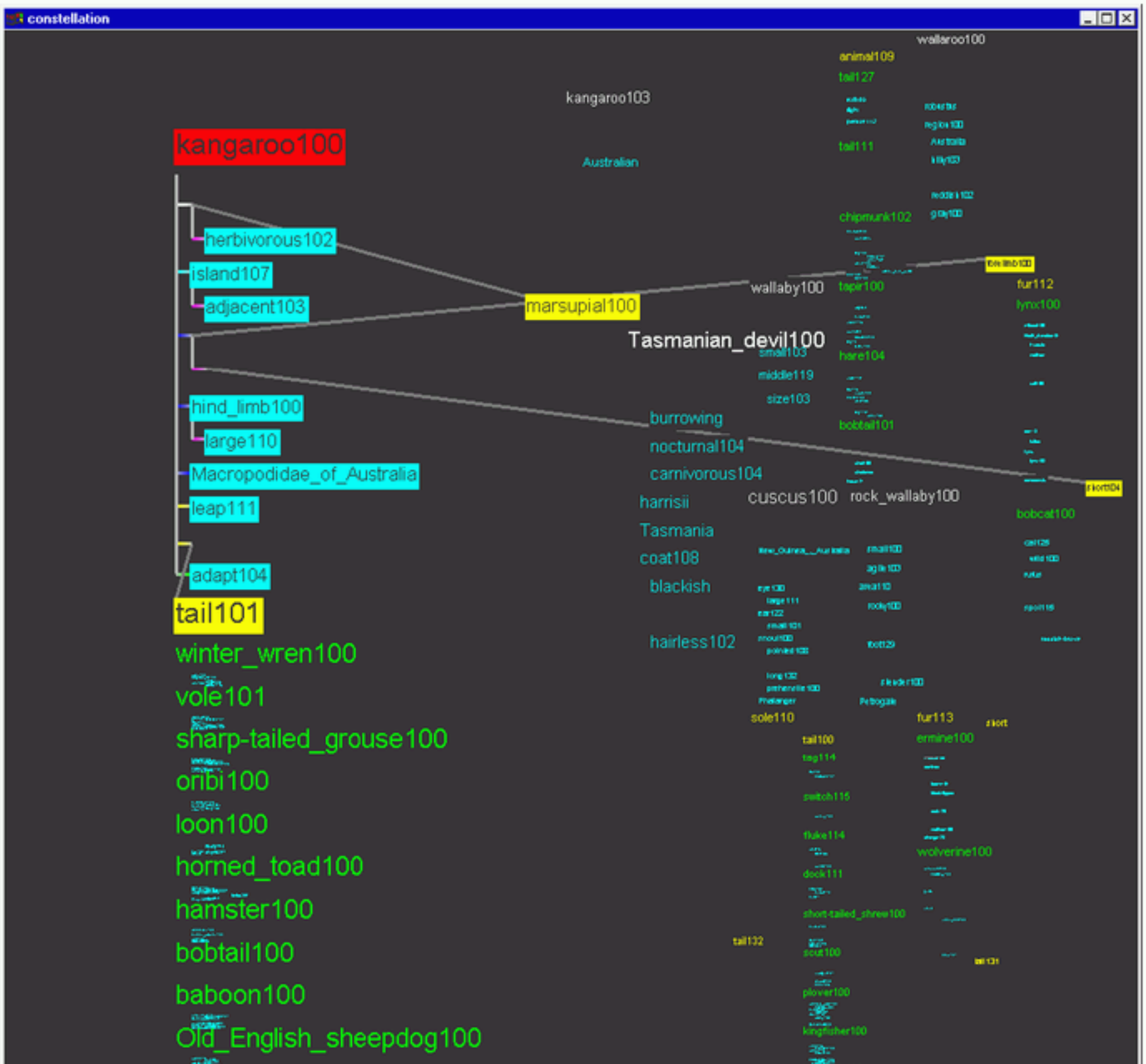
- change salience instead of toggle drawing

why? closed world assumption

- implicit assumption: if not visible, doesn't exist
- easy to forget previous actions

  draw false negative conclusions

Single vs. Multiple Word Instances

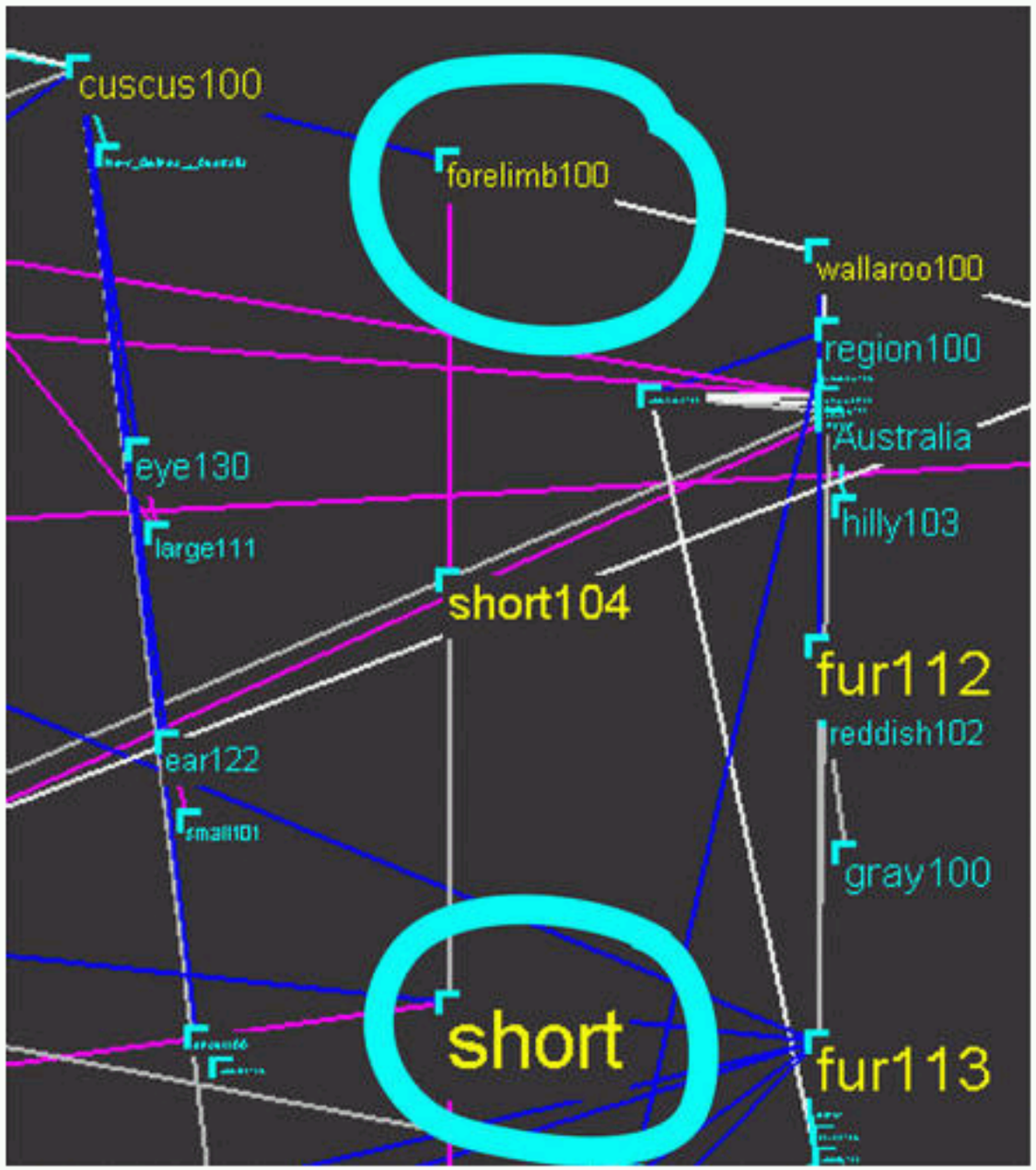


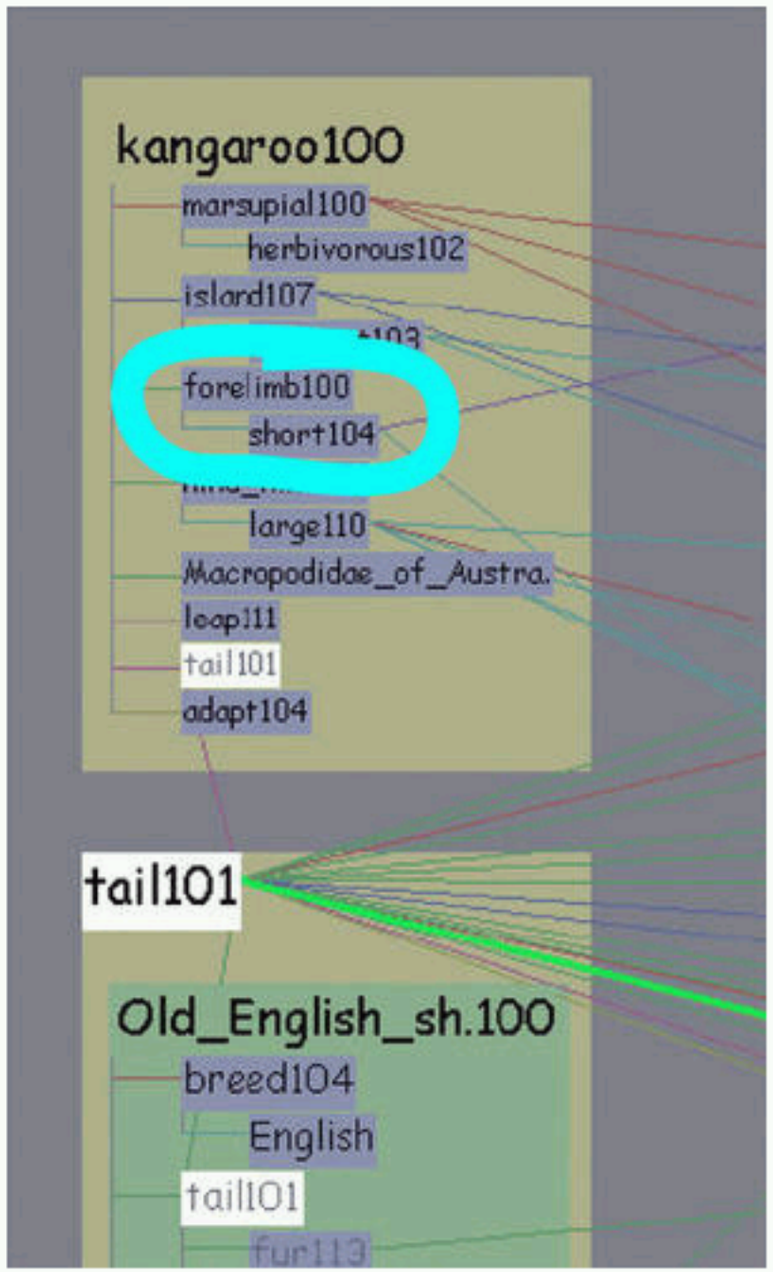
Information Density

early prototype: poor

Information Density

design tradeoff with visual salience

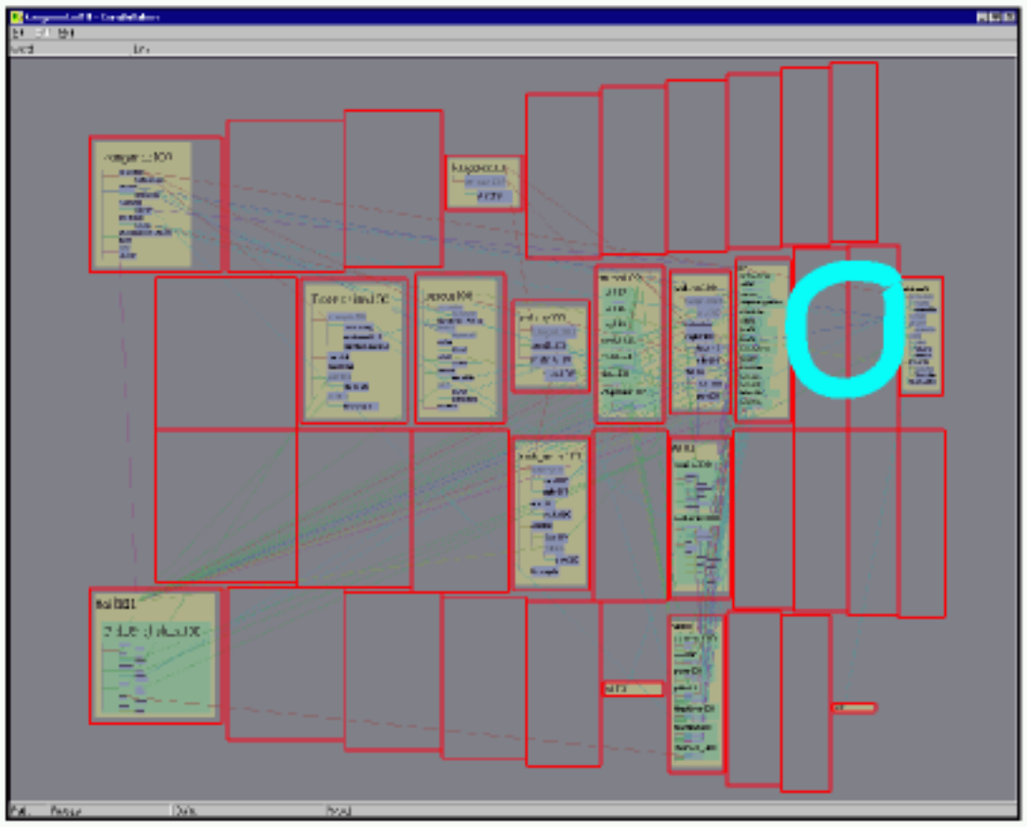


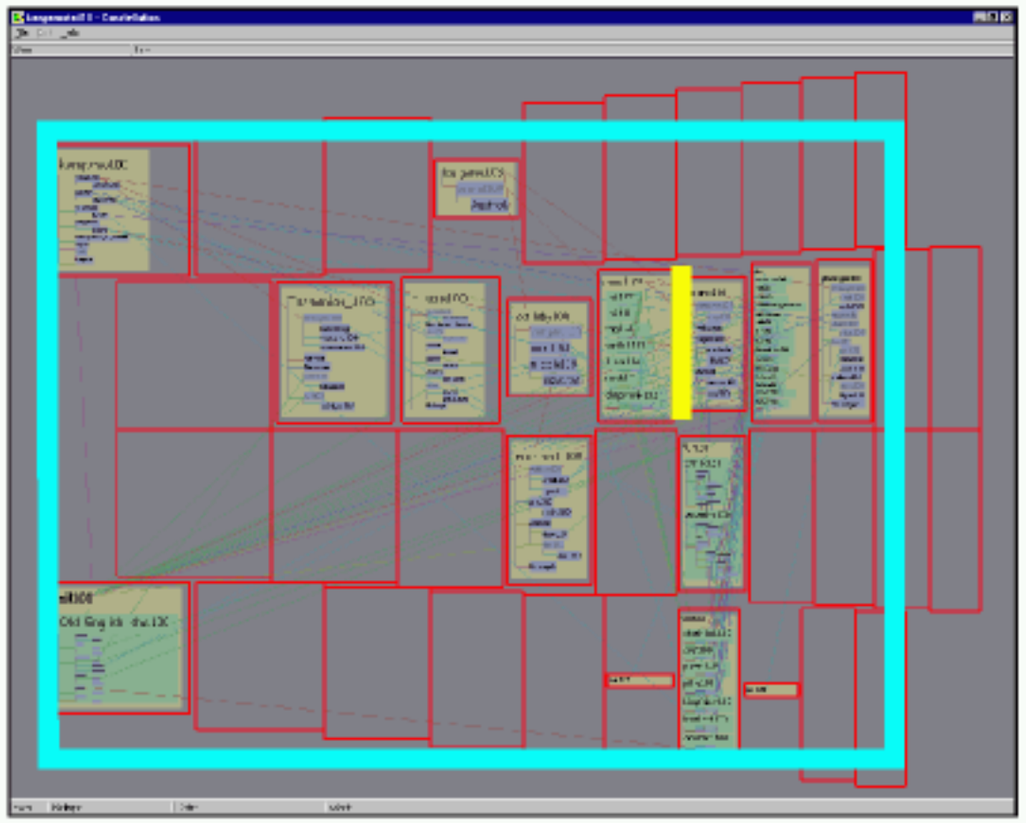


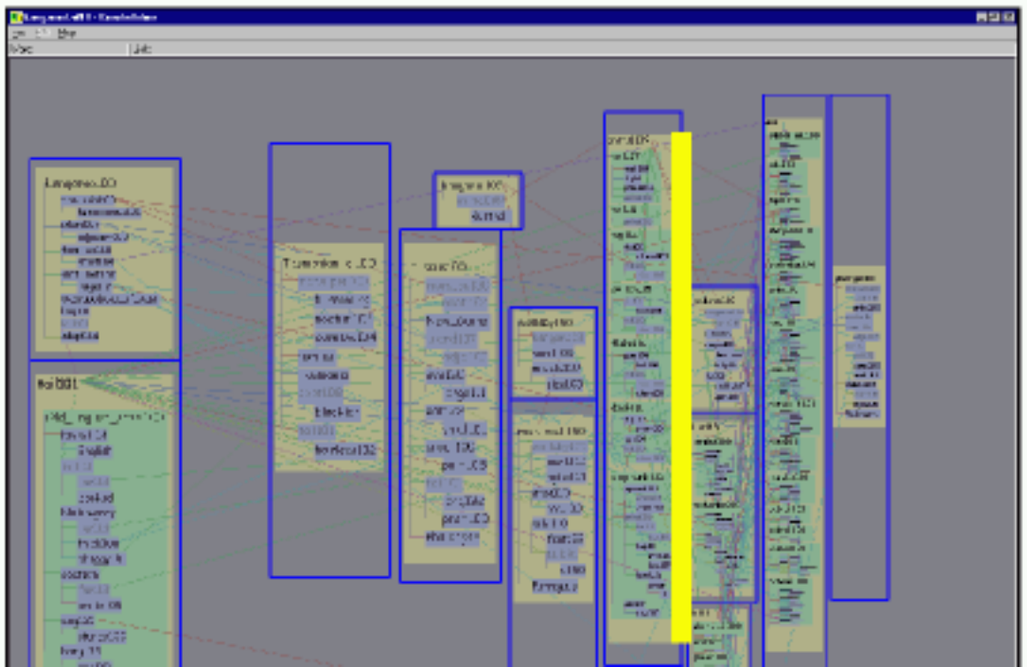
[graphics.stanford.edu/papers/munzner_thesis/html/node11.html#noncanonfig]

Information Density

grid adjustment

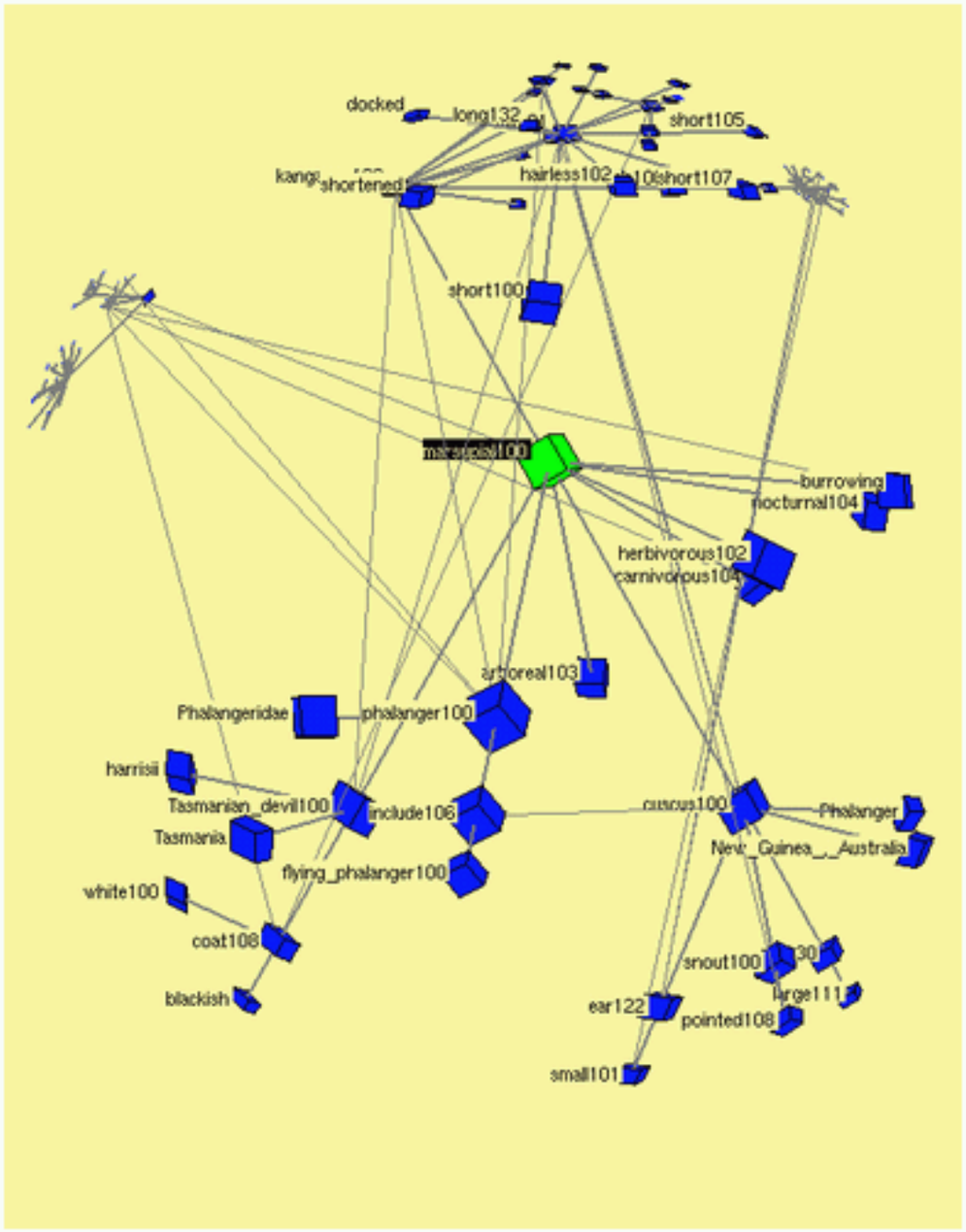


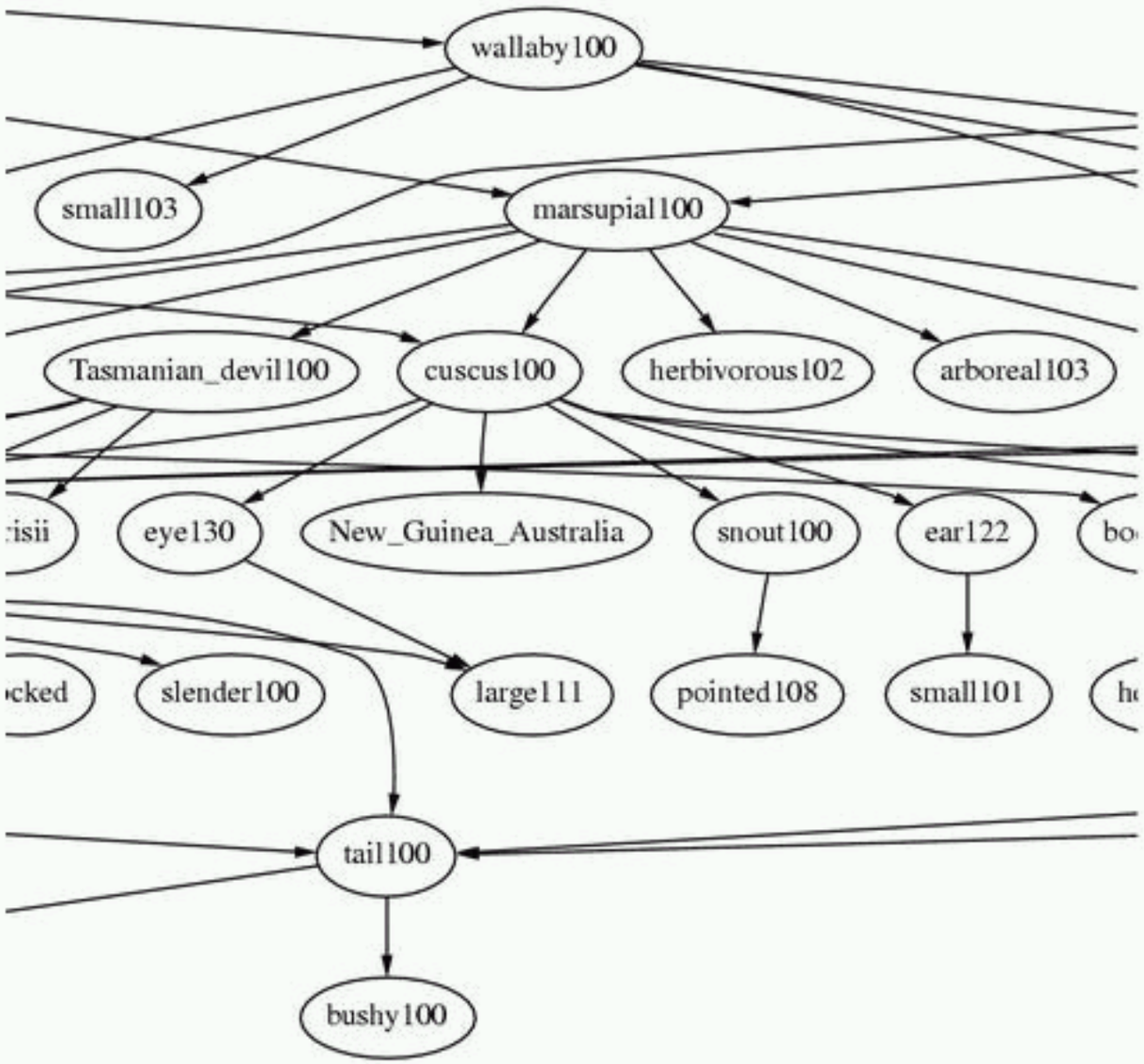




Task-oriented design

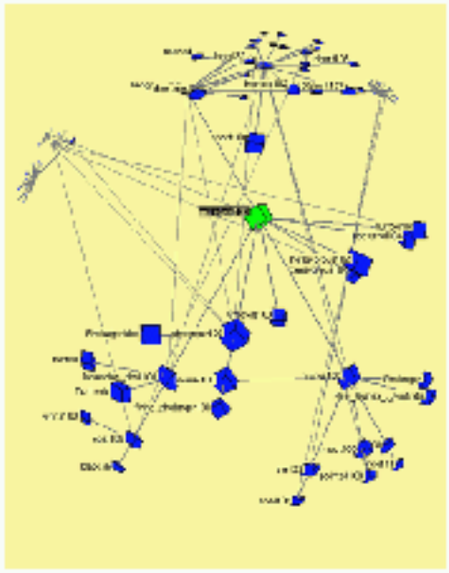
previous methods

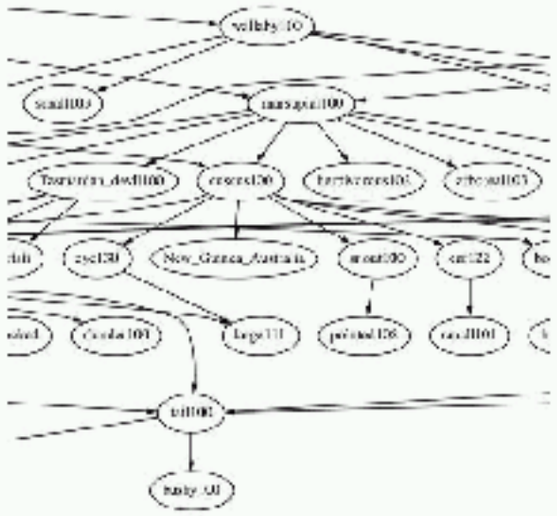




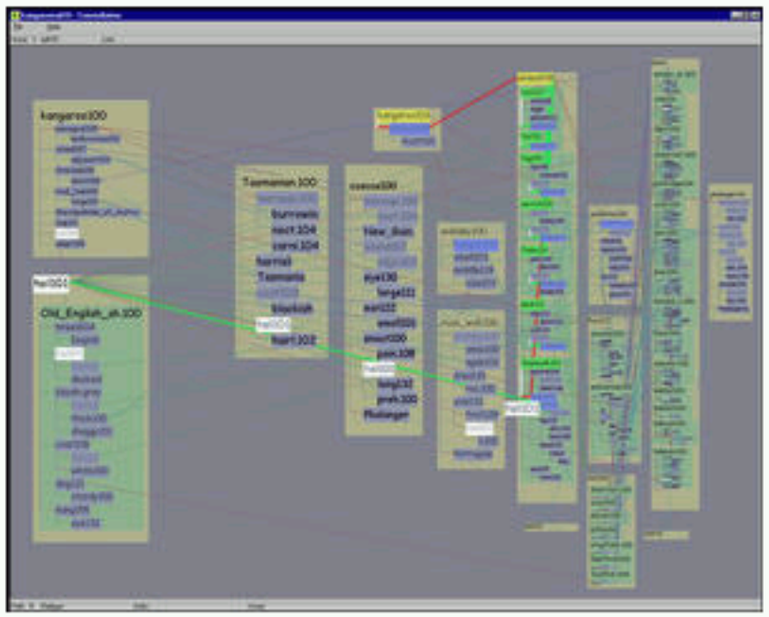
[graphics.stanford.edu/papers/munzner_thesis/html/node10.html#dotconstfig]

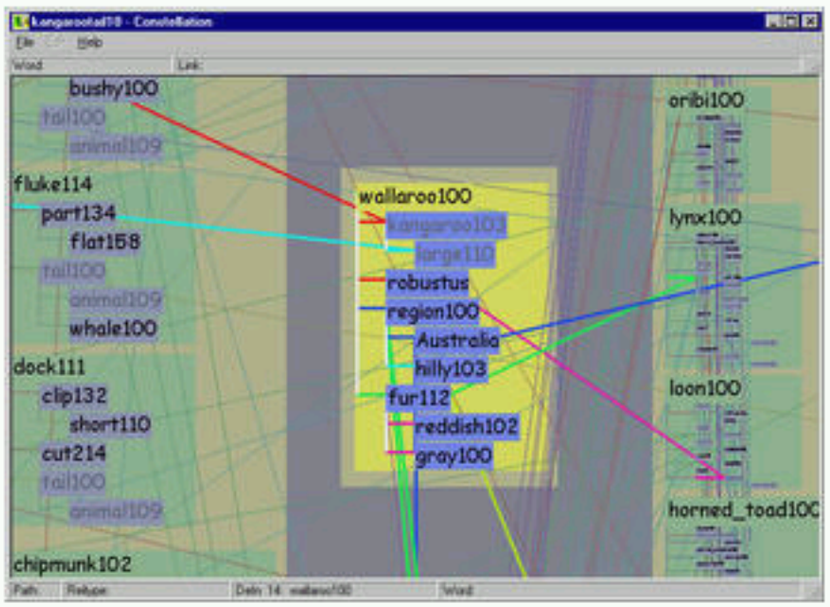
Task-oriented design





task-specific methods





[graphics.stanford.edu/papers/munzner_thesis/html/node10.html#layoutefffig]