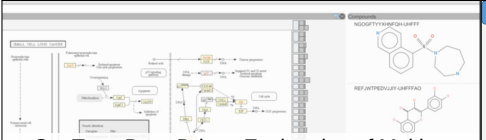


ConTour: Data-Driven Exploration of Multi-Relational Datasets for Drug Discovery

Christian Partl, Alexander Lex, Marc Streit, Hendrik Strobelt, Anne-MaiWassermann, Hanspeter Pfister and Dieter Schmalstieg

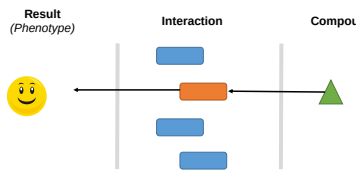
Domain Problem

UNDERSTANDING DRUG DISCOVERY

Scenario 1:
Targeted interaction, understood mechanism, desired outcome

= biological receptor
= biological target
= chemical compound
= potential target
= direct target
= inhibits target

Result *(Phenotype)* — Interaction — Compound



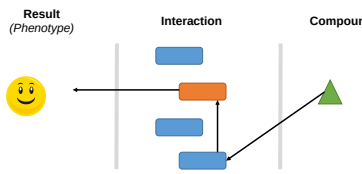
Domain Problem

UNDERSTANDING DRUG DISCOVERY

Scenario 2:
Indirect interaction, understood mechanism, desired outcome

= biological receptor
= biological target
= chemical compound
= potential target
= direct target
= inhibits target

Result *(Phenotype)* — Interaction — Compound



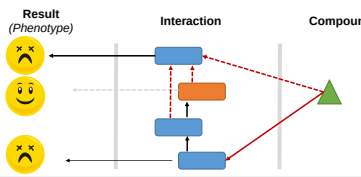
Domain Problem

UNDERSTANDING DRUG DISCOVERY

Scenario 3:
Complex interactions, mechanism poorly understood, multiple outcomes

= biological receptor
= biological target
= chemical compound
= potential target
= direct target
= inhibits target

Result *(Phenotype)* — Interaction — Compound

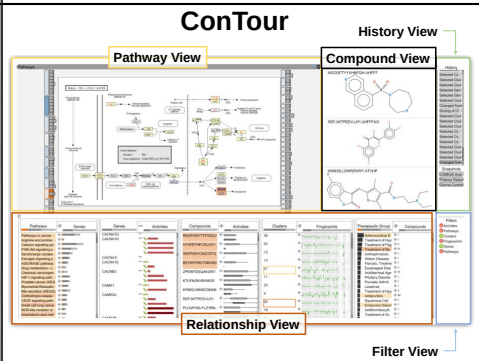


Domain Problem

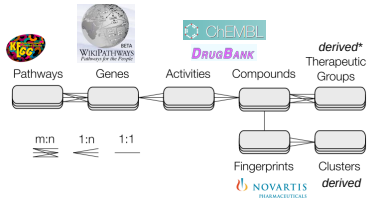
Drug Discovery Main Goals

- Identify a drug's mechanism of action
- Identify the biological process a drug modulates
- Identify new drugs for specific therapeutic indications

ConTour

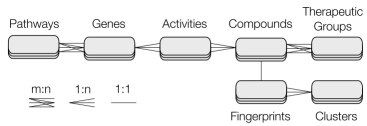


Data Abstraction



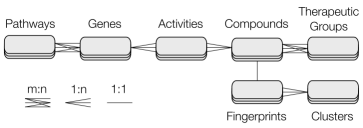
* Derived using a scheme propose by the Prous Integrity database

Data Abstraction



"The drug discovery domain problem can be **generalized** to the problem of analysing ***multi-relational*** datasets […] Consequently, we argue that our approach is **applicable to many other problems**."

Data Abstraction



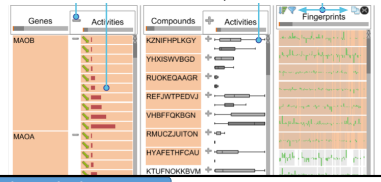
"The multi-relational data exploration problem can be interpreted as a **graph exploration problem** where each item of each dataset represents a node and the relationships between the items are the edges"

Task Analysis

- **T1**: Identify Related Items

Item selection and highlighting

Clicking, not hovering, on an item also moves all related items in columns to the top



Task Analysis

- **T1**: Identify Related Items

Selection-based filters

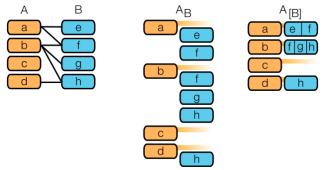
Filter choices when multiple items are selected

Task Analysis

- **T1**: Identify Related Items

Nesting

Simple Nesting

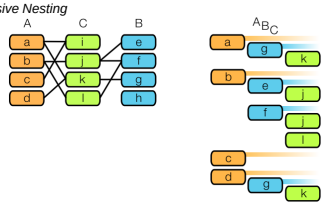


Task Analysis

- **T1**: Identify Related Items

Nesting

Recursive Nesting

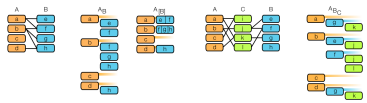


Task Analysis

- **T2:** Identify Items that Share a Relationships with a Set of Items

Nesting

Simple Nesting — *Recursive Nesting*

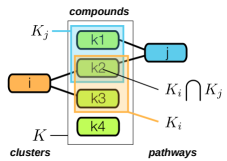


Task Analysis

- **T3:** Analyse Network Enrichment

Enrichment Score

Judging how specific two items are when compared to a third



$$s_{i,j}(K) = \frac{|K_i \cap K_j| / |K_j|}{|K_i| / |K|}$$

Where:
I = clusters
K = compounds
J = Pathways
S(i,j) = pair score

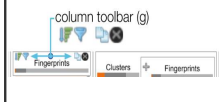
*I assume they take care of divide by 0?

Task Analysis

- **T4:** Rank Items

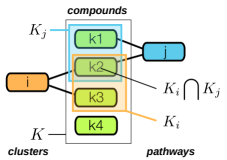
Sorting by interest

Sort alpha-numerically



Enrichment Score

Sort by enrichment score

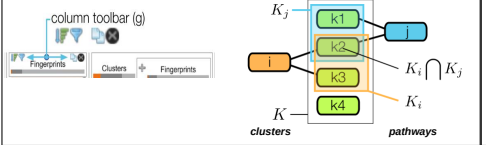


Task Analysis

- **T4:** Rank Items

Sorting by interest
Sort alpha-numerically

Enrichment Score
Sort by enrichment score



Task Analysis

- **T5:** Filter Items
  Depends on tasks 1 and 2

Navigation

Local Filter : filter within a specific column

Global Filter: remove items that are not connected to the source column

Task Analysis

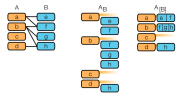
- **T5:** Filter Items
  Depends on tasks 1 and 2

Navigation

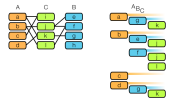
Selection-based filters

Nesting

Simple Nesting



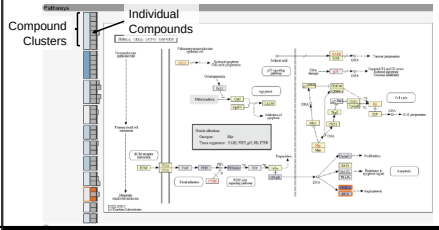
Recursive Nesting



Task Analysis

- **T6**: View items in detail

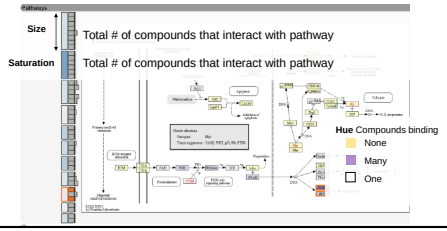
Pathway View
Marks



Task Analysis

- **T6**: View items in detail

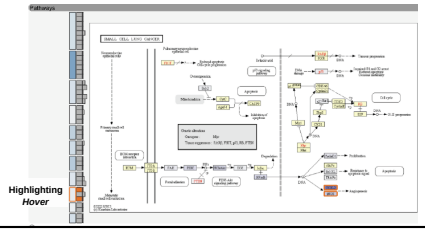
Pathway View
Channels



Task Analysis

- **T6**: View items in detail

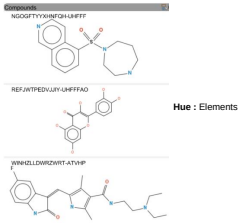
Pathway View
Linked Views



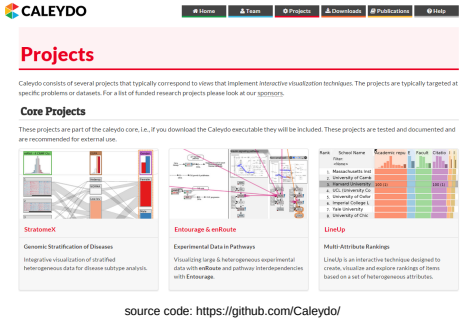
Task Analysis

- **T6**: View items in detail

Compound View

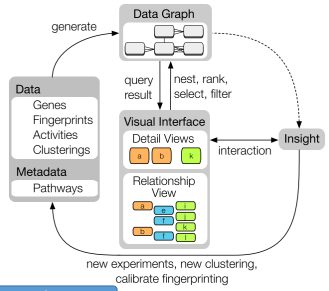


Implementation Details

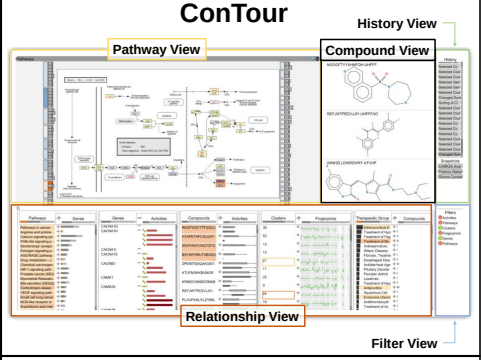


source code: https://github.com/Caleydo/

ConTour



ConTour



Relationship View

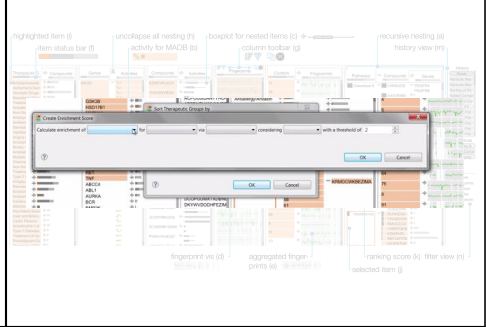
- **There are facets within the relationship view**
- **Combination of tabular data and plots**

Relationship View

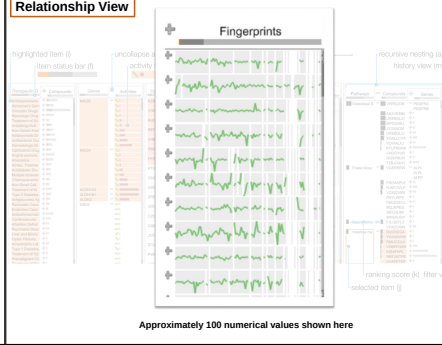
Algorithm Design

Relationship View

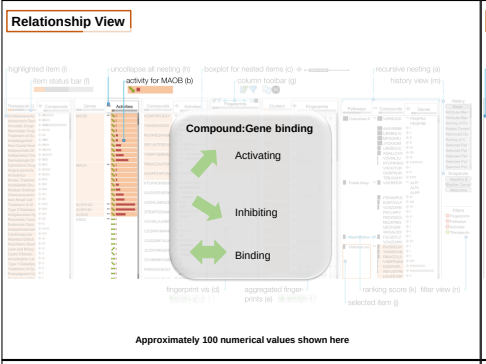
Relationship View

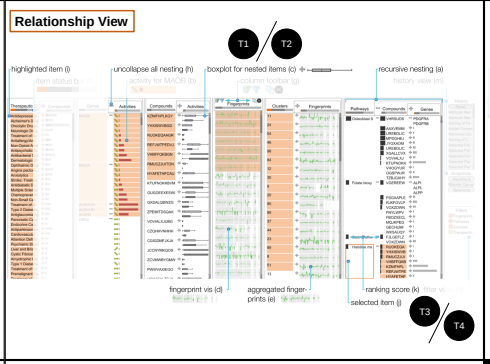


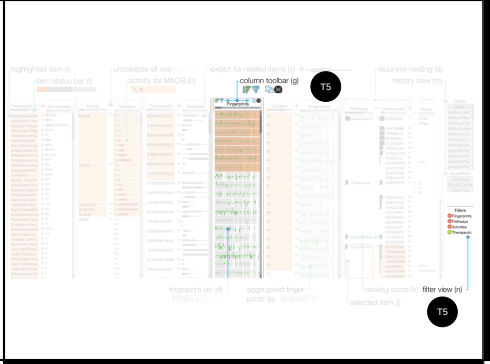
Relationship View



Approximately 100 numerical values shown here







Conclusions

System	ConTour
What : Data	Multi-relational databases; node-link graph; clusters *(derived)*
Why : Tasks	Discovery; drill down; highlight relationships
How: Multiple Views	Relationship view; pathway view; compound view; history and filters
How: Facet	Side-by-side linked views, containing tabular data, bar plots, glyphs
How: Selection & Highlighting	Linked highlighting across facets; automatic sorting
How: Filtering	Drag and drop (nesting); user control (navigation)
How: Ranking & sorting	Enrichment score; highlight; user control (navigation)
How: Encode	Simple marks with manipulation of hue and saturation (pathway view)
Scale:	Dozens of columns; upper limit on HD display appears to about 20. Thousands of data items. Up to 8 simultaneous views for compounds; only 1 for pathways

Concluding Thoughts

- Seems like a very good tool for use on structured datasets
- When there are indirect (inferred) relationships, it would be good to highlight this with some uncertainty
- What about incomplete relationships?

Backup