

Monte Carlo Tree Search for Generating Interactive Data Analysis Interfaces



Background: Interfaces and Tasks

General purpose interfaces:

Can do a lot, but complex to use.



Specialized (Precision) interfaces

Designed for specific task, but easy to use, simple



Problem: Building Interfaces is Hard 🛞

1. Expensive to learn user's analysis tasks + build interface

2. Dashboard builders limited in analysis complexity or require programming

3. Prior work auto-generates forms using database contents, but ignores analysis task may be complex if DB has many tables+attrs

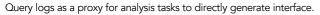


Tasks

High costs \rightarrow interfaces mainly built for high profile tasks that are "worth it"

"Unpopular" tasks ignored, but they are important for their users!

Previous Work: Precision interface V1[1]





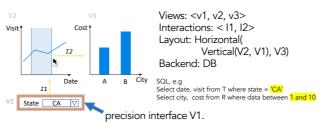
Yiru Chen Columbia University.

Main Contribution

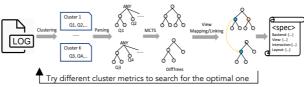
Compared to previous work which only generates a set of widgets, this work enhances the generated interface quality from:

- Automatically generate an optimal set of <u>views</u>, including <u>widgets and visualizations</u>, and the interaction over the views.
- 2. Consider hierarchical layout as well as the usability in terms of how easy to express the query log.
- 3. A generic frontend engine which takes a specification as input and output a web app.

What is an Interface?



Overview



Problem Statement: Given query log, view types, Find lowest cost interface and layout where all queries can be expressed.

1. Cluster metrics

- Project attribute union compatible
- Overall query similarity etc.

2. Parse: Queries modeled as abstract syntax trees

3. Monte Carlo tree Search:

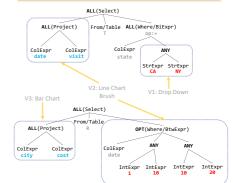
- The layout and the selected views are intertwined with the process of identifying subtree differences between the input query ASTs. Search space is very large.
- Method: UCT algorithm
- Reward: Randomly assign views and greedily link k times to pick the lowest cost.
- 4. Cost Estimation for view mapping and linking:

appropriateness of each view + usefulness + layout score

Example: View Mapping and Linking

After we get the final difftree, we search for the optimal view mapping and linking. Layout will be expressed by assigning layout node: Horizontal and Vertical to ALL node.

SELECT date, visit FROM T WHERE state='CA'; SELECT date, visit FROM T WHERE state='NV'; SELECT city, cost FROM R WHERE date BETWEEN 1 AND 10; SELECT city, cost FROM R WHERE date BETWEEN 10 AND 20;



Generic Frontend Engine



Layout sensitive interface generated from 10 Sloan Digital Sky Survey (SDSS 2017) gueres.



Interactive visualization interface generated from falcon flights query log.

