BOP: Modular Platform for Learned Query Optimizer Research

Gautam Mittal, Zongheng Yang

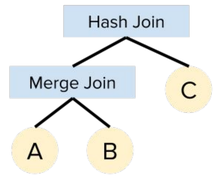
riselab UC Berkeley

Query → Execution Plan

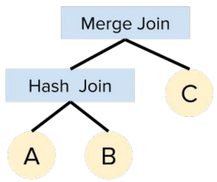
Query

```
SELECT MIN(chn.name), MIN(t.title)
FROM char_name AS chn, title AS t
WHERE t.production_year > 1990
  AND t.id = 4
  AND chn.id = 9;
```

Plan

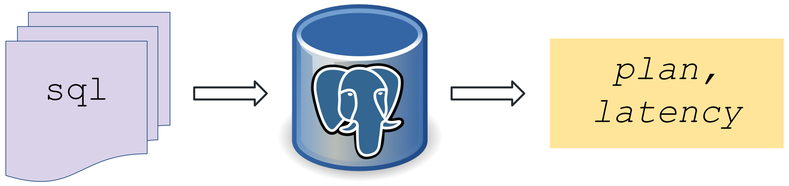



```
SELECT MAX(k.keyword), n.name
FROM keyword AS k, name AS n
WHERE k.keyword IN ('superhero',
                    'sequel')
  AND n.name LIKE '%Downey%Robert%'
```



Query execution plan space is exponentially large and often relies on human-engineered cost models to generate the best plan.

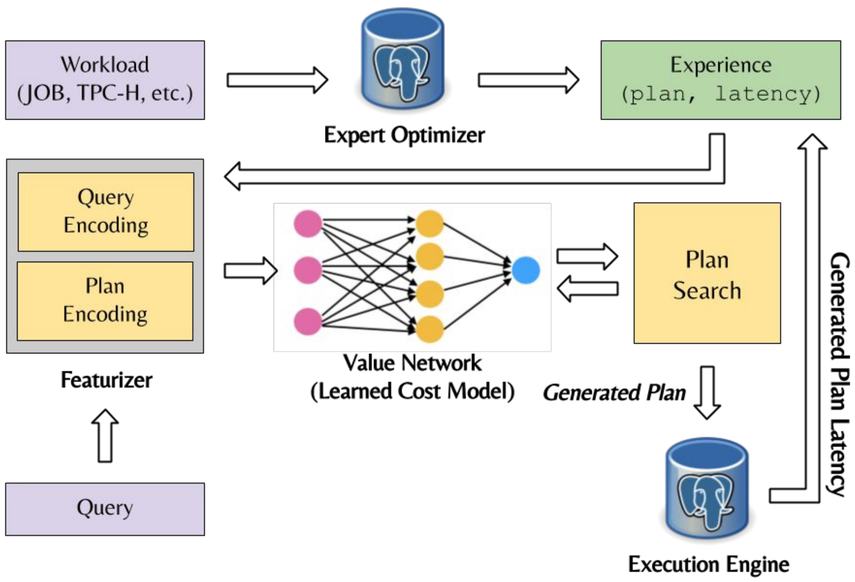
Expert Bootstrapping



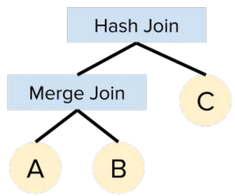
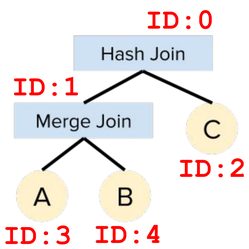
Deep reinforcement learning agent learns an initial policy π_0 from a human-engineered (expert) model.

Why is this necessary?	The cost of executing poor plans can result in queries that take hours to complete.
Value Iteration	Use π_0 to learn an optimal policy through retraining agent based on generated plans + latency.

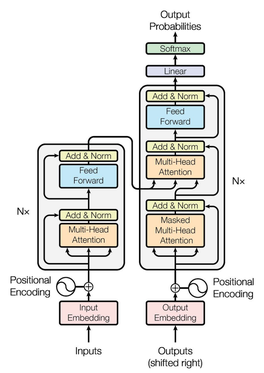
System Architecture



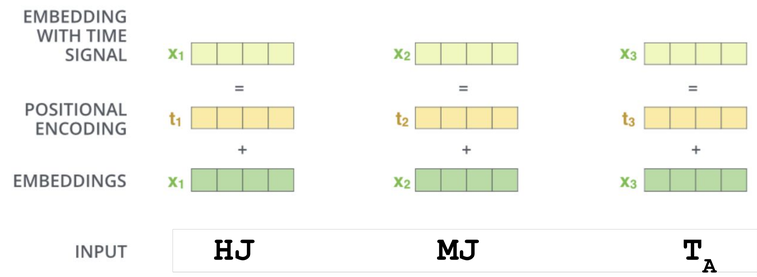
Encoding Plans as Sequences

Pre-order Seq. Encoding	 `[HJ, MJ, T_A, T_B, T_C]`
Parent Positional Embeddings (parent's ID)	 `[-1, 0, 1, 1, 0]`

Sequence encodings are processed by a Transformer model.

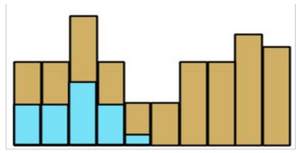


Vaswani et al. 2017. Attention is all you need. In Advances in Neural Information Processing Systems, pages 6000–6010.

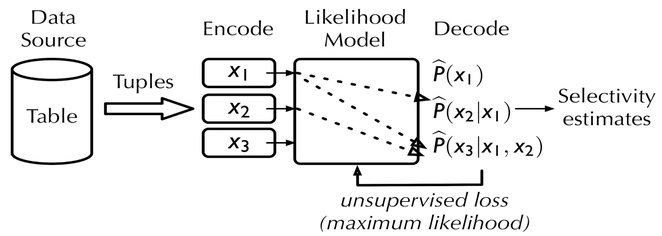


Query Encoding

Flexibility of system allows for novel improvements to featurization and cost model.

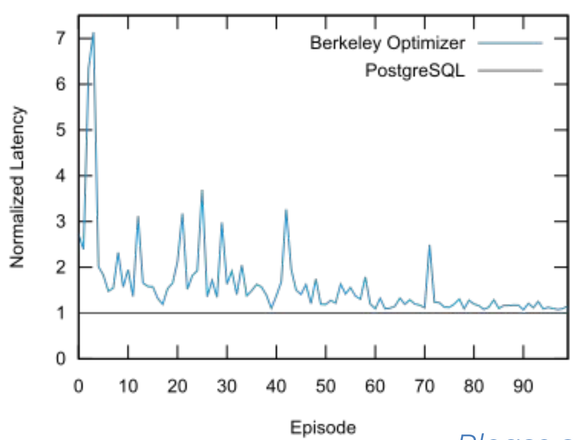


Histogram-based Selectivity Estimation



Naru: *Deep Unsupervised Cardinality Estimation*

Evaluation



Workload: JOB (70 queries)

Model: Transformer

Featurization: Positional embeddings + pre-order encoding + hist. estimation

Please send us feedback:
gbm@berkeley.edu, zongheng@cs.berkeley.edu