

A quick taste:

Explaining outputs in modern data analytics

[VLDB2016]

Replace “Datalog” with Differential Dataflow:

1. More expressive (e.g. Map/Reduce/Aggr),
2. Incrementally maintained.

Explanations:

For a target subset of the observed output,
identify a subset of the input such that:
compute yields output containing the subset.

Differential Dataflow

current conclusions (ongoing)

Functional queries + timestamping = hugely powerful:

1. Automatic incrementalization (even w/loops).
2. Throughput scale-out w/o concurrency. (*c.f.* Calvin)
3. Shared state between multiple queries. (*c.f.* RDBMS)

Recovering RBDMS features w/o sacrificing usability.
We are sacrificing e.g. transactions. This is not OLTP.

A quick taste:

Explaining outputs in modern data analytics
[VLDB2016]

Replace “Datalog” with Differential Dataflow:

1. More expressive (e.g. Map/Reduce/Aggr),
2. Incrementally maintained.

Explanations:

For a target subset of the observed output,
identify a subset of the input such that:
compute yields output containing the subset.