Logging an Egg: Datalog on E-Graphs

Philip Zucker

Draper Laboratory

June 13, 2022

Datalog: Databases, Logic, and Proofs.

$$\forall x, y. edge(x, y) \implies path(x, y)$$

$$\frac{edge(x, y)}{path(x, y)}$$

$$path(X, Y) := edge(X, Y).$$

$$path(X, Y) := path(X, Z), edge(Z, Y).$$

Search right hand side in database. Insert left side. Repeat.

Datalog: Applications

- Graph Problems
- Worklist Algorithms
- You can program in it. Explicit control flow
- Mutually Recursive Analyses
- Program Analysis
 - ► Doop ¹
 - ▶ DDisasm ²



¹https://bitbucket.org/yanniss/doop/src/master/

Datalog vs Prolog

- ► Pattern matching vs Unification
- ► Top down vs Bottom up
- ► Complete vs Incomplete Search
- ► Tabling

E-Graphs

- Datastructure for terms and equalities
- ► Egg ³: Efficient Rust library
- ► Term Rewriting $?a + 0 \rightarrow ?a$
- ► An E-Graph Never Forgets
- Maximize sharing up and down.
- Rule ordering
- ightharpoonup ?a + (?b + ?c) = (?a + ?b) + ?c. ?a + -?a = 0
- x + (-x + 10)



Applications of E-Graphs

- SMT and other Theorem Proving
- ► Compiler Optimization. PEG ⁴
- ► Herbie ⁵
- Query Optimization
- Szalinski CAD
- ▶ YOGO ⁶

⁴https://rosstate.org/publications/eqsat/

⁵https://herbie.uwplse.org/

⁶https://www.jameskoppel.com/files/papers/yogo-preprint.pdf 📱

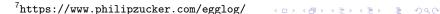
Egglog0⁷

- E-Graphs are a Database
- The database holds terms and equality relation
- Supports ordinary datalog with terms
- Pattern variables bind to eclasses
- Rules: query using RHS (e-matching multipattern), instantiate and insert LHS
- Special equality = is E-graph Equality

$$add(Y,X) = E :- add(X,Y) = E.$$

Queries e-match and return all results.

$$?- add(succ(zero), succ(Y)) = Z.$$





Egg Multipatterns

- ► Upstreamed to egg⁸.
- Multipatterns vs Guards
- ► Threads e-matching compiler env binding between patterns.

Demo

Example: Injectivity

- $ightharpoonup \forall a, b. f(a) = f(b) \implies a = b$
- Example: Constructors, Negation, constant addition
- Unification

$$X = Y$$
, $Xs = Ys :- cons(X, Xs) = cons(Y, Ys)$. $X = Y :- X + Z = Y + Z$.

Example: Memory Simplification⁹

- ► Alias Analysis + Simplification
- ► SMTlib theory of arrays
- Many SMT theories are expressible as Horn Clauses (side conditions)

```
\label{eq:continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous_continuous
```

⁹https://www.philipzucker.com/egglog/?example=mem.pl= → < ≥ → へ ?

Example: Equation Solving

- Isolation
- Extract terms without variables

$$\mathsf{sub}\,(\mathsf{Z}\,,\mathsf{X}) \,=\, \mathsf{Y} \,:-\,\,\mathsf{add}\,(\mathsf{X}\,,\mathsf{Y}) \,=\, \mathsf{Z}$$

Example: Reflection

- ► Hypothetical reasoning
- ► Boolean algebraic reasoning

$$\begin{array}{lll} A = B : - & true = eq(A,B). \\ true = eq(A,B) : - & A = B. \end{array}$$

Example: Uniqueness Quantification¹⁰

- ► Common in universal constructions in category theory
- Skolemize existentials $\forall x, P(x) \implies \exists y, Q(x, y)$ becomes $\forall x, P(x) \implies Q(x, f(x))$
- ▶ Uniqueness Property $\forall a, b.P(a) \land P(b) \implies a = b$

¹⁰https://www.philipzucker.com/egglog/?example=cat1.pl > 2 > 2 > 2 <

Related Work

- ► Relational E-Matching ¹¹
- SMT Multipatterns
- ► Souffle Egg ¹²
- Egg-lite

//www.hytradboi.com/2022/writing-part-of-a-compiler-in-datalog

¹¹https://arxiv.org/abs/2108.02290

¹²https:

Questions?

- Thanks to Yihong Zhang, Remy Yisu Wang, Max Willsey, Zachary Tatlock, Alessandro Cheli, Cody Roux, James Fairbanks, and Evan Patterson for their helpful discussions.
- ► DARPA Grant