A generic approach to proximity-based matching

Maximilian Donnermair

FEEDBACK

- Titel vllt anpassen, wobei „generality“ eh an sich passt
- Symbol vs constant vs term vs function
- Semantics of variables; deltas are set of variables, syntactically
- Dv „function“ erklären!
- Sinn hinter deltas erklären!

Problem statement

- Quantitative Matching
  - Given λ-threshold and proximity degrees between ground terms
  - Goal: $M; \emptyset; 1 \Rightarrow \emptyset; S; \delta$ with $\delta \geq \lambda$
  - M: set of problems, e.g.: $f(x, x) \preccurlyeq g(a, b)$, S: $\{ x \mapsto ? \}$
- Proximity calculations with T-Norms (Triangular Norms)
  - $\otimes$ instead of $\wedge$
- Current matching algorithm: only with Minimum-T-Norm
  - $x \otimes y = \min(x, y)$

- Goal: **generalizing inference rules**

„Standard“ Inference rules

- **Decompose**: transform $\{f(t_1, \ldots, t_n) \preccurlyeq g(s_1, \ldots, s_n)\}$ into $\{ t_i \preccurlyeq s_i | \, 1 \leq i \leq n\}$
  - Proximity of functions $R(f, g)$ gets „added“ to final proximity degree
  - Clash rule (exception): different arity, proximity too low
- **Solve**: $\{x \preccurlyeq t\} \Rightarrow \begin{Bmatrix} x \mapsto t_0 \\ x \approx pc(t) \end{Bmatrix}$ (Branching vs Compact representation)
  - $pc(t) \coloneqq \{(s, \alpha) | R(t, s) \geq \alpha\}$
  - Clash rule (exception): proximity class of some (sub)term is empty
- **Merge**: transform $\{x \approx pc(t), x \approx pc(s)\}$ into $\{x \approx pc(t) \cap pc(s)\}$
  - Intersection of proximity classes
  - Clash rule (exception): t and s don‘t have same tree structure

Example

- $M \coloneqq \{f(x,x) \preccurlyeq_{\lambda=0.5} g(a,b)\}$ with $a \approx_{0.8} c, b \approx_{0.9} c, a \approx_{0.2} b, f \approx_{0.7} g$
  - Decompose into $\{x \preccurlyeq a, x \preccurlyeq b\}$
  - „Branching“ method: Solve one, instantiate, then decompose
  - „Compact“ method: Solve both, then merge
- Approximation degree decreases over time

- New implicit approach: λ-threshold increases
  - We get a set of degree constraints for the variables

New inference rules

Dec: Decomposition

$$\{f(t_1,\ldots,t_n) \preceq_\delta g(s_1,\ldots,s_m)\} \uplus M;\ S;\ \Delta \otimes \delta \geq \lambda \Longrightarrow$$
$$M \cup \{t_i \preceq_{\delta_i} s_i \mid 1 \leq i \leq n\};\ S;\ \Delta \otimes (\textstyle\bigotimes_{i=1}^{n} \delta_i) \geq Res_\otimes(\mathcal{R}(f,g),\lambda)$$

if $n = m$ and where $n \geq 0$; and $\delta_1,\ldots,\delta_n$ are fresh degree variables.

Sol: Solve

$$\{x \preceq_\delta t\} \uplus M;\ S;\ \Delta \otimes \delta \geq \lambda \Longrightarrow$$
$$M;\ S \cup \{x \approx \mathbf{ext}^{\delta}_{\mathcal{R},\lambda}(t) =: \mathbf{t}^\delta\};\ \Delta \otimes (\textstyle\bigotimes_{p\in Pos(t)} dv(\mathbf{t}^\delta, p)) \geq \lambda$$

Mer: Merge

$$M;\ S \uplus \{x \approx \mathbf{ext}^{\delta_t}_{\mathcal{R},\lambda}(t) =: \mathbf{t}^{\delta_t}, x \approx \mathbf{ext}^{\delta_s}_{\mathcal{R},\lambda}(s) =: \mathbf{s}^{\delta_s}\};$$
$$\Delta \otimes (\textstyle\bigotimes_{p\in Pos(t)} dv(\mathbf{t}^{\delta_t}, p)) \otimes (\textstyle\bigotimes_{p\in Pos(s)} dv(\mathbf{s}^{\delta_s}, p)) \Longrightarrow$$
$$M;\ S \cup \{x \approx \mathbf{t}^{\delta_t} \sqcap \mathbf{s}^{\delta_s}\};\ \Delta \otimes (\textstyle\bigotimes_{p\in Pos} \gamma_p^x \geq \lambda$$

where $Pos = Pos(t) = Pos(s)$ and $\gamma_p^x := dv(\mathbf{t}^{\delta_t}, p) \otimes dv(\mathbf{s}^{\delta_s}, p))$.

Term extensions

- Because proximity classes are not compact enough
- $|pc(f(a,b))| = |pc(f)| * |pc(a)| * |pc(b)|$
  - $\{f(a,b), f(a,c), f(a,a), f(b,a), f(b,b), \dots, g(a,b), g(a,c), \dots\}$
- $|ext(f(a,b))| = |pc(f)| + |pc(a)| + |pc(b)|$
  - $\{f,g\}(\{a,b,c\},\{a,b,c\})$
- Here without approximation degrees
- Intersection of classes position-wise

Example

- $\{f\left(f(x,y), g\big(y,a,h(x)\big)\right) \preccurlyeq_\delta h(f(b,g(b)), h(f(c),b,g(c)))\}$
  - Decompose several times
- $\{x \preccurlyeq b, y \preccurlyeq g(b), y \preccurlyeq f(c), a \preccurlyeq b, x \preccurlyeq c\}$
  - Solve (4x)
- $\{x = ext(b) = \{(a, 0.95), (b, 1), (c, 0.75)\},$
  $x = ext(c) = \{(a, 0.85), (b, 75), (c, 1)\},$
  $y = ext\big(g(b)\big) = \{(f, 0.7), (g, 1), (h, 0.8)\}(\{(a, 0.95), (b, 1), (c, 0.75)\}),$
  $y = ext\big(f(c)\big) = \{(f, 1), (g, 0.7), (h, 0.9)\}(\{(a, 0.85), (b, 0.75), (c, 1)\})\}$
  - Merge twice
- $\{x = \{(a, 0.95 \otimes 0.85), (b, 1 \otimes 0.75), (c, 0.75 \otimes 1)\},$
  $y = \{(f, 0.7 \otimes 1), (g, 1 \otimes 0.7), (h, 0.8 \otimes 0.9)\}$
  $(\{(a, 0.95 \otimes 0.85), (b, 1 \otimes 0.75), (c, 0.75 \otimes 1)\})\}$

Conclusion

- Sound-/Completeness proofs will determine completeness of Clash rules
- Improving/Simplifying constraint management
    - „worst case“ method
    - eliminating unsatisfiable instantiations early on
- Anti-unification
- „Fully fuzzy“ signatures