When Almost Is Not Even Close: Remarks on the Approximability of HDTP

Tarek R. Besold Robert Robere

AI Research Group, Institute of Cognitive Science, University of Osnabrück

University of Toronto

01. August 2013

Outline

Parameterized Complexity Theory (1)

Parameterized Decision Problem

An instance of a parameterized decision problem $\mathcal{P}$ is a tuple (x, k), where $x \in \{0, 1\}^*$ is a string describing the problem and $k \in \mathbb{Z}$, which is called the parameter of the problem (codifying other aspects of the problem besides n).

Fixed Parameter Tractability (FPT)

A parameterized decision problem $\mathcal{P}$ is fixed parameter tractable, written $\mathcal{P} \in$ FPT, if it is solvable in time bounded by $f(k) \cdot |x|^{O(1)}$, where $f(k)$ is some computable function of the parameters and $|x|^{O(1)}$ denotes a polynomial of the length of the input.

As suggested by notation, let FPT denote the class of all fixed parameter tractable problems.

Parameterized Complexity Theory (2)

W[1]-membership & W[1]-hardness

W[1] is the class of problems solvable by constant depth combinatorial circuits with at most 1 gate with unbounded fan-in on any path from an input gate to an output gate.
A parameterized problem $\mathcal{P}$ is W[1]-hard if every problem in $W[1]$ is reducible to $\mathcal{P}$ under a parameterized reduction.

W[1] $\neq$ FPT **can be seen as analogous to** P $\neq$ NP **in classical complexity.**[1]

[1] FPT $=$ W[0] and W[i] $\subseteq$ W[j] for all $i \leq j$ (most likely: W[i] $\subset$ W[j]).

Approximation Theory

Approximability Classes

During the course of this presentation, let...

- ...PTAS denote the class of all NP optimization problems that admit a polynomial-time approximation scheme.
- ...APX be the class of NP optimization problems allowing for constant-factor approximation algorithms.
- ...APX-*poly* be the class of NP-optimization problems allowing for polynomial-factor approximation algorithms.

Please note that PTAS $\subseteq$ APX $\subseteq$ APX-*poly* (with each inclusion being proper in case P $\neq$ NP).

Outline

1. ~~Parameterized Complexity & Approximation Theory~~ ✓
2. Heuristic-Driven Theory Projection (HDTP)
3. Complex Analogies: A Parameterized Analysis of HDTP
4. Almost Right: An Approximability Analysis of HDTP
5. Conclusion and Future Work

Heuristic-Driven Theory Projection (1)

Heuristic-Driven Theory Projection (HDTP)

- Computing **analogical relations and inferences** (domains given as many-sorted first-order logic representation/many-sorted term algebras) using a generalization-based approach.
- Base and target of analogy defined in terms of **axiomatisations**, i.e., given by a finite set of formulae.
- Aligning pairs of formulae by means of **anti-unification** (extending classical Plotkin-style first-order anti-unification to a restricted form of higher-order anti-unification).
- Proof-of-concept applications in modeling mathematical reasoning and concept blending in mathematics.

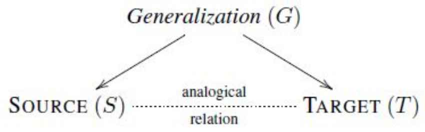


Figure : Analogy-making in HDTP.

Heuristic-Driven Theory Projection (3)

Anti-Unification

- Dual to the unification problem (see, e.g., logic programming or automated theorem proving).
- **Generalizing terms** in a meaningful way, yielding for each term an **anti-instance** (distinct subterms replaced by variables).
- Goal: Finding the **most specific anti-unifier**.
- Plotkin: For a proper definition of generalization, for a given pair of terms there always is exactly one least general generalization (up to renaming of variables).
- Problem: Structural commonalities embedded in different contexts possibly not accessible by first-order anti-unification.

Heuristic-Driven Theory Projection (4)

Restricted Higher-Order Anti-Unification

- First-order terms extended by **introducing variables taking arguments** (first-order variables become variables with arity 0), making a term either a first-order or a higher-order term.
- Class of **substitutions restricted** to (compositions of) the following four cases:

1. Renamings ρ^{F,F^*}: $F(t_1,\dots,t_n) \xrightarrow{\rho^{F,F^*}} F^*(t_1,\dots,t_n)$.
2. Fixations ϕ^F_c: $F(t_1,\dots,t_n) \xrightarrow{\phi^F_f} f(t_1,\dots,t_n)$.
3. Argument insertions $\iota^{F,F^*}_{G,i}$:
$$F(t_1,\dots,t_n) \xrightarrow{\iota^{F,F^*}_{G,i}} F^*(t_1,\dots,t_i,G(t_{i+1},\dots,t_{i+k}),t_{i+k+1},\dots,t_n).$$
4. Permutations π^{F,F^*}_α: $F(t_1,\dots,t_n) \xrightarrow{\pi^{F,F^*}_\alpha} F^*(t_{\alpha(1)},\dots,t_{\alpha(n)})$.

Heuristic-Driven Theory Projection (5)

Examples of higher-order anti-unifications:

$$\begin{tikzcd}[column sep=tiny] & f(X) \arrow[dl, "\rho^X_Y"'] \arrow[dr, "\rho^X_Z"] & \\ f(Y) & & f(Z) \end{tikzcd}$$

(a)

$$\begin{tikzcd}[column sep=tiny] & f(X) \arrow[dl, "\phi^X_a"'] \arrow[dr, "\phi^X_b"] & \\ f(a) & & f(b) \end{tikzcd}$$

(b)

$$\begin{tikzcd}[column sep=tiny] & F(a) \arrow[dl, "\phi^F_f"'] \arrow[dr, "\phi^F_g"] & \\ f(a) & & g(a) \end{tikzcd}$$

(c)

$$\begin{tikzcd}[column sep=tiny] & F(a,b,c) \arrow[dl, "\iota^{F,F'}_{X,2}"'] \arrow[dr, "\iota^{F,F''}_{G,1}"] & \\ F'(a,b,X,c) & & F''(a,G(b,c)) \end{tikzcd}$$

(d)

$$\begin{tikzcd}[column sep=tiny] & F(a,b) \arrow[dl] \arrow[dr, "\pi^{F,F'}_{\alpha}"] & \\ F(a,b) & & F'(b,a) \end{tikzcd}$$

(e)

Heuristic-Driven Theory Projection (6)

Solar System	Rutherford Atom
sorts $real,\ object,\ time$ **entities** $sun : object,\ planet : object$ **functions** $mass : object \to real \times \{kg\}$ $dist : object \times object \times time \to real \times \{m\}$ $force : object \times object \times time \to real \times \{N\}$ $gravity : object \times object \times time \to real \times \{N\}$ $centrifugal : object \times object \times time \to real \times \{N\}$ **predicates** $revolves\_around : object \times object$ **facts** $\alpha_1 : mass(sun) > mass(planet)$ $\alpha_2 : mass(planet) > 0$ $\alpha_3 : \forall t : time : gravity(planet, sun, t) > 0$ $\alpha_4 : \forall t : time : dist(planet, sun, t) > 0$ **laws** $\alpha_5 : \forall t : time,\ o_1 : object,\ o_2 : object :$ $dist(o_1, o_2, t) > 0 \wedge gravity(o_1, o_2, t) > 0$ $\to centrifugal(o_1, o_2, t) = -gravity(o_1, o_2, t)$ $\alpha_6 : \forall t : time,\ o_1 : object,\ o_2 : object :$ $0 < mass(o_1) < mass(o_2) \wedge dist(o_1, o_2, t) > 0 \wedge$ $centrifugal(o_1, o_2, t) < 0$ $\to revolves\_around(o_1, o_2)$	**sorts** $real,\ object,\ time$ **entities** $nucleus : object,\ electron : object$ **functions** $mass : object \to real \times \{kg\}$ $dist : object \times object \times time \to real \times \{m\}$ $coulomb : object \times object \times time \to real \times \{N\}$ **facts** $\beta_1 : mass(nucleus) > mass(electron)$ $\beta_2 : mass(electron) > 0$ $\beta_3 : \forall t : time : coulomb(electron, nucleus, t) > 0$ $\beta_4 : \forall t : time : dist(electron, nucleus, t) > 0$

Heuristic-Driven Theory Projection (7)

types
$real,\ object,\ time$
constants
$X : object,\ Y : object$
functions
$mass : object \longrightarrow real \times \{kg\}$
$dist : object \times object \times time \rightarrow real \times \{m\}$
$F : object \times object \times time \rightarrow real \times \{N\}$
$centrifugal : object \times object \times time \longrightarrow real \times \{N\}$
predicates
$revolves\_around : object \times object \times object$
facts
$\gamma_1 : mass(X) > mass(Y)$
$\gamma_2 : mass(Y) > 0$
$\gamma_3 : \forall t : time : F(X, Y, t) > 0$
$\gamma_4 : \forall t : time : dist(X, Y, t) > 0$
laws
$\gamma_5* : \forall t : time,\ o_1 : object,\ o_2 : object :$
$dist(o_1, o_2, t) > 0 \wedge F(o_1, o_2, t) > 0$
$\rightarrow centrifugal(o_1, o_2, t) = -F(o_1, o_2, t)$
$\gamma_6* : \forall t : time,\ o_1 : object,\ o_2 : object :$
$0 < mass(o_1) < mass(o_2) \wedge dist(o_1, o_2, t) > 0 \wedge\ centrifugal(o_1, o_2, t) < 0$
$\rightarrow revolves\_around(o_1, o_2)$

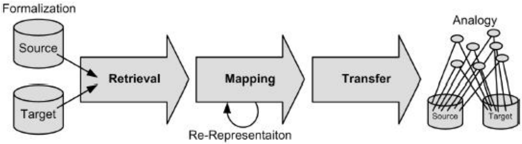
Outline

1. ~~Parameterized Complexity & Approximation Theory~~ ✓
2. ~~Heuristic-Driven Theory Projection (HDTP)~~ ✓
3. Complex Analogies: A Parameterized Analysis of HDTP
4. Almost Right: An Approximability Analysis of HDTP
5. Conclusion and Future Work

Complexity of HDTP (1)

HDTP is naturally split into two mechanisms:

- Analogical matching of input theories.
- Re-representation of input theories by deduction in FOL.



⇒ Re-representation is undecidable (undecidability of FOL).

⇒ Focus on mechanism for analogical matching.

Complexity of HDTP (2)

Problem 1. F Anti-Unification
Input: Two terms f, g, and a natural $k \in \mathbb{N}$
Problem: Is there an anti-unifier h, containing at least k variables, using only **renamings and fixations**?

Problem 2. FP Anti-Unification
Input: Two terms f, g, and naturals $l, m, p \in \mathbb{N}$.
Problem: Is there an anti-unifier h, containing at least l 0-ary variables and at least m higher arity variables, and two substitutions σ, τ using only renamings, fixations, and **at most** p **permutations** such that $h \xrightarrow{\sigma} f$ and $h \xrightarrow{\tau} g$?

Problem 3. FPA Anti-Unification
Input: Two terms f, g and naturals $l, m, p, a \in \mathbb{N}$.
Problem: Is there an anti-unifier h, containing at least l 0-ary variables, at least m higher arity variables, and two substitutions σ, τ using renamings, fixations, at most p permutations, and **at most** a **argument insertions** such that $h \xrightarrow{\sigma} f$ and $h \xrightarrow{\tau} g$?

Complexity of HDTP (3)

Complexity of HDTP (Higher-Order Anti-Unification)

1. **F Anti-Unification is solvable in polynomial time**.
2. Let m be the maximum number of higher arity variables and p be the maximum number of permutations applied. Then **FP Anti-Unification is** NP**-complete and** W[1]**-hard w.r.t. parameter set** $\{m, p\}$.
3. Let r be the maximum arity and s be the maximum number of subterms of the input terms. Then **FP Anti-Unification is in** FPT **w.r.t. parameter set** $\{s, r, p\}$.
4. **FPA Anti-Unification is** NP**-complete and** W[1]**-hard w.r.t. parameter set** $\{m, p, a\}$.

Outline

1. ~~Parameterized Complexity & Approximation Theory~~ ✓
2. ~~Heuristic-Driven Theory Projection (HDTP)~~ ✓
3. ~~Complex Analogies: A Parameterized Analysis of HDTP~~ ✓
4. Almost Right: An Approximability Analysis of HDTP
5. Conclusion and Future Work

Approximability Analysis (1)

- FP Anti-Unification W[1]-hard to compute for parameter set m, p (m number of higher-arity variables, p number of permutations).
  $\Rightarrow$ No polynomial-time algorithm computing "sufficiently complex" generalizations (i.e., with lower bound on number of higher-arity variables), upper bounding number of permutations (W[1]-hardness for single permutation).
- **What if one considers generalizations which merely approximate the "optimal" generalization in some sense?**

Approximability Analysis (2)

Complexity of a Substitution

The complexity of a basic substitution σ is defined as

$$C(s) = \begin{cases} 0, & \text{if } \sigma \text{ is a renaming.} \\ 1, & \text{if } \sigma \text{ is a fixation or permutation.} \\ k+1, & \text{if } \sigma \text{ is a } k\text{-ary argument insertion.} \end{cases}$$

The complexity of a restricted substitution $\sigma = \sigma_1 \circ \cdots \circ \sigma_n$ (i.e., the composition of any sequence of unit substitutions) is the sum of the composed substitutions: $C(\sigma) = \sum_{i=1}^{n} C(\sigma_i)$.

Approximability Analysis (3)

Consider problem of finding generalization which **maximizes complexity over all generalizations**:

- Complex generalization would contain "most information" present over all of the generalizations chosen (i.e., maximizing the "information load").
- Using approximability results on MAXCLIQUE:

Approximation Complexity of HDTP Analogy-Making

FP anti-unification is not in APX **(i.e., is hard for** APX**-poly).**

Outline

1. ~~Parameterized Complexity & Approximation Theory~~ ✓
2. ~~Heuristic-Driven Theory Projection (HDTP)~~ ✓
3. ~~Complex Analogies: A Parameterized Analysis of HDTP~~ ✓
4. ~~Almost Right: An Approximability Analysis of HDTP~~✓
5. Conclusion and Future Work

Conclusion and Future Work

- Analogy-making in HDTP is widely not tractable or approximable!
- Taking the Tractable AGI Thesis (**see talk tomorrow afternoon!**) into account, the suitability of HDTP as basis for a general model for high-level cognitive capacities or a general cognitive architecture seems questionable.
- Main question(s) for future research:
  - How can the computation of generalizations via restricted higher-order anti-unification be constrained in a meaningful way as to remain polynomially-solvable?
  - Similarly: How can the underlying KR formalism be restrained/modified?
  - How can the undecidability of the re-representation mechanism directly be addressed (and mitigated)?

Contact the authors: **tbesold@uos.de** or **robere@cs.toronto.edu**.