Completeness in the Second Level of the Polynomial Time Hierarchy through Syntactic Properties

Edwin Pin

Universidad Central de Venezuela,
Facultad de Ciencias, Departamento de Matemática,
edwin.pin@ciens.ucv.ve

Abstract. In the mid-90's Neil Immerman and José Medina initiated the search for syntactic tools to prove **NP**-completeness. They conjectured that if a problem defined by the conjunction of an Existential Second Order sentence and a First Order Formula is **NP**-complete then the Existential Second Order formula defines an **NP**-complete problem. This property was called *superfluity* of First Order Logic with respect to **NP**. Few years later it was proved by Nerio Borges and Blai Bonet that superfluity is true for the universal fragment of First Order Logic and with respect not only to **NP** but for a major rank of complexity classes. Our work extends that result to the Second Level of the Polynomial-Time Hierarchy [9].

1 Background

We consider the descriptive approach to our objects of study. Although this is not the way it is commonly done, we introduce complexity classes syntactically. Other concepts, as reducibility and completeness, will be understood in the context of Descriptive Complexity.

Let $\mathcal{L}$ be a logical language closed under disjunctions and closed under conjunctions with first-order formulas. The complexity class $\mathbf{C}$ captured by $\mathcal{L}$ is the set of all decision problems defined by sentences in $\mathcal{L}$ i.e.

$$\mathbf{C} := \{\mathrm{MOD}[\varphi] : \varphi \in \mathcal{L}\},$$

where $\mathrm{MOD}[\varphi]$ is the set of all finite structures that satisfy sentence φ.

This notion of complexity classes follows from [3], in which $\mathbf{C}$ is asked to be nice, closed under finite unions and also dependent on a family of proper complexity functions [8].

Let SO_k be the language of second order (SO) sentences with at most k alternations of quantifiers, starting with an existential one. So, for every natural number k, SO_k consists of all second order sentences with the prenex form

$$\exists \mathbf{R}_1 \forall \mathbf{R}_2 \ldots \mathcal{Q}_k \mathbf{R}_k \varphi,$$

where $\mathcal{Q}_k$ is existential if k is odd and universal if k is even, $\mathbf{R}_j$ is a tuple of relation variables and φ is a first order (FO) sentence. Notice that by Fagin's Theorem $\mathbf{NP} = \{\mathrm{MOD}[\varphi] : \varphi \in \mathrm{SO}_1\}$ [5].

We pay special attention to the case $k = 2$ i.e. the language SO_2 of sentences of the form $\exists \mathbf{R}_1 \forall \mathbf{R}_2 \varphi$, and also to the complexity classes

$$\Sigma_2^p := \{\mathrm{MOD}[\varphi] : \varphi \in \mathrm{SO}_2\}$$
$$\Pi_2^p := \{\mathrm{MOD}[\varphi] : \neg\varphi \in \mathrm{SO}_2\}.$$

Another important computational concept is *reducibility.* Informally, a problem A is reducible to another problem B if there is an easily computable map f from instances of A to instances of B such that $x \in A \iff f(x) \in B$. The function f is called a *reduction.* The type of reductions we need are *first order projections* (or fops, to simplify). For a general treatment of their properties we refer the reader to [1]. If A is reducible to B (via fops) we write $A \leq_{fop} B$.

Our idea of completeness in a complexity class depends on our notion of reduction. A problem B is *hard* (via fops) in the complexity class $\mathbf{C}$ or $\mathbf{C}$-hard if $A \leq_{fop} B$ for every problem $A \in \mathbf{C}$. We say that B is $\mathbf{C}$-complete (via fops) if $B \in \mathbf{C}$ and it is $\mathbf{C}$-hard.

Theorem 1. *The following problems are Σ_2^p-complete:*

1. QSat$_2$
2. QUnsat$_2$
3. $\exists\exists!$Sat
4. 2CC

The description of all these problems appears in the appendix at the end of this document and an analysis of their completeness can be check on [9].

2 Further Concepts

Suppose Ψ is a sentence in SO_1. According to [7] a first order sentence φ is *superfluous* if there exists a fop ρ from SAT to $\mathrm{MOD}[\Psi \wedge \varphi]$ such that $\rho(\mathcal{A}) \models \varphi$ for every structure $\mathcal{A}$ representing a CNF Boolean formula. An immediate consequence of this definition is the following proposition:

Proposition 1. *[7] If the conjunction $\Psi \wedge \varphi$ defines an **NP**-complete with Ψ a sentence in SO_1 and φ a superfluous sentence in FO then $MOD[\Psi]$ is **NP**-complete.*

It is known that the expressive power of first-order logic is strictly less than the expressive power of existential second-order. It is reasonable then to conjecture that the hypothesis of φ being superfluous in Proposition 1 is not necessary:

Conjecture 1. [7] If the conjunction $\Psi \wedge \varphi$ defines an **NP**-complete with Ψ a sentence in SO_1 and φ a sentence in FO then $\mathrm{MOD}[\Psi]$ is **NP**-complete.

There is a partial answer to Conjecture 1 in [2].

Proposition 2. *[2] If the conjunction $\Psi \wedge \varphi$ defines an **NP**-complete problem with Ψ a sentence in SO_1 and φ an universal sentence in FO then $MOD[\Psi]$ is **NP**-complete.*

2.1 Superfluity, Consistency and Universality

The notion of superfluity is generalized to any complexity class in the following definitions:

Definition 1. *Let σ and τ be two vocabularies, $\mathcal{L}$ a logic, $\mathbf{C}$ the complexity class captured by $\mathcal{L}$, $\mathcal{L}'$ a fragment of $\mathcal{L}$ and $\rho : STRUC[\sigma] \to STRUC[\tau]$ a fop.*

1. *A sentence $\varphi \in \mathcal{L}'[\tau]$ is superfluous with respect to ρ if $\rho(\mathcal{A}) \models \varphi$ for every $\mathcal{A} \in STRUC[\sigma]$.*
2. *$\varphi \in \mathcal{L}'$ is superfluous with respect to $\mathcal{L}$ if for every sentence $\Psi \in \mathcal{L}$, the $\mathbf{C}$-completeness of $MOD[\Psi \wedge \varphi]$ implies the $\mathbf{C}$-completeness of $MOD[\Psi]$.*
3. *$\mathcal{L}'$ is superfluous with respect to $\mathcal{L}$ (or $\mathbf{C}$) if every sentence $\varphi \in \mathcal{L}'$ is superfluous with respect to $\mathcal{L}$.*

Medina's conjecture can be paraphrased as FO is superfluous with respect to **NP**. To established the results of ∀FO (the set of universal first order sentences) as a superfluous logic we need to introduce the notion of consistency of formulas and universality of problems.

Definition 2. *Let $\varphi(\boldsymbol{x})$ be a formula in $FO[\sigma]$, n be a natural number, and $\boldsymbol{u} \in n^k$, where k is the length of the first-order-variable tuple $\boldsymbol{x}$. We say that $\langle \varphi(\boldsymbol{x}), \boldsymbol{u} \rangle$ is n-consistent if there is a σ-structure $\mathcal{A}$ such that $||\mathcal{A}|| = n$ and $\mathcal{A} \models \varphi(\boldsymbol{u})$. If $S \subseteq STRUC[\sigma]$, we say that $\langle \varphi(\boldsymbol{x}), \boldsymbol{u} \rangle$ is n-consistent in S if there is a structure $\mathcal{A} \in S$ such that $||\mathcal{A}|| = n$ and $\mathcal{A} \models \varphi(\boldsymbol{u})$. If there is no risk of confusion, we abbreviate by just saying that $\varphi(\boldsymbol{u})$ is n-consistent (in S).*

Definition 3. *Let's suppose now that $\sigma = \langle R_1^{a_1}, \ldots, R_r^{a_r}, c_1, \ldots, c_s \rangle$ and S the same as before. Let n and t be two natural numbers.*

1. *S is $(n, 0)$-universal if for every $m \geq n$ and every sequence $b_1, \ldots, b_s \in m$ there is a structure $\mathcal{A} \in S$ with $\|\mathcal{A}\| = m$ and such that $\mathcal{A} \models (c_1 = b_1) \wedge \cdots \wedge (c_s = b_s)$.*
2. *S is (n, t)-universal if for every $m \geq n$, every sequence of σ-literals $L_1, \ldots, L_t$ (that is, $L_j(\boldsymbol{x})$ is equal to $R_{i_j}(\boldsymbol{x})$ or $\neg R_{i_j}(\boldsymbol{x})$), every sequence of tuples $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_t$ with $\boldsymbol{u}_j \in m^{a_{i_j}}$ and every sequence $b_1, \ldots, b_s \in m$, the m-consistency of*

$$\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_t, b_1, \ldots, b_s) \equiv \bigwedge L_j(\boldsymbol{u}_j) \wedge \bigwedge (c_k = b_k) \tag{1}$$

implies its m-consistency in S (that is, if there are models of (1) of cardinality m, at least one belongs to S).

Definition 4. *A family $\mathcal{F}$ of problems over a vocabulary σ is complete and universal for a complexity class* **C** *if*

1. *every problem in $\mathcal{F}$ is* **C**-*complete;*
2. *There is a sequence $\{n_k\}_{k\geq 0}$ and a natural number m such that for every $k \geq m$ there is a (n_k, k)-universal problem S_{n_k} in $\mathcal{F}$ that contains all the σ-structures $\mathcal{A}$ with $||\mathcal{A}|| < n_k$.*

Theorem 2. *[2] Let* **C** *be a complexity class captured by $\mathcal{L}$ with $\mathrm{FO} \subseteq \mathcal{L}$. If* **C** *contains a complete and universal family $\mathcal{F}$, then $\forall\mathrm{FO}$ is superfluous with respect to* **C**.

We use the last theorem to prove superfluity is valid in the Second Level of the Polynomial-Time Hierarchy.

3 Main Results

To prove that $\forall\mathrm{FO}$ is superfluous with respect to Σ_2^p we need to generate an appropriate complete and universal family. We make use of the problem 2CC and the following propositions.

Definition 5. *If S is a problem over a vocabulary σ, we define for each $n \in \mathbb{N}$:*

$$S_n := S \cup \{\mathcal{A} \in \mathrm{STRUC}[\sigma] : ||\mathcal{A}|| < n\} \tag{2}$$

and the family of problems

$$\mathcal{F}(S) := \{S_n\}_{n\geq 2}. \tag{3}$$

Lemma 1. 2CC *is $(2k+1, k)$-universal for every $k \geq 1$.*

Corollary 1. *Given any natural number $n \geq 2$, $2\mathrm{CC}_n$ is $(2k+1, k)$-universal for every $k \geq 1$.*

Lemma 2. *For every natural number $n \geq 2$ the problem $2\mathrm{CC}_n$ belongs to Σ_2^p.*

Lemma 3. *For every natural number $n \geq 2$ the problem $2\mathrm{CC}_n$ is Σ_2^p-hard.*

Theorem 3. *$\mathcal{F}(2\mathrm{CC})$ is a complete and universal family in Σ_2^p.*

Theorem 4. *$\forall\mathrm{FO}$ is superfluous with respect to Σ_2^p.*

The proofs of all these theorems can be check on [9]. We have prove also that superfluity is valid in the dual class Π_2^p.

Lemma 4. *$(2\mathrm{CC})^c$ is $(2k+5, k)$-universal for every $k \geq 1$.*

Theorem 5. *$\mathcal{F}((2\mathrm{CC})^c)$ is a complete and universal family in Π_2^p.*

Theorem 6. *$\forall\mathrm{FO}$ is superfluous with respect to Π_2^p.*

4 Conclusions

Theorem 2 is used in [2] to prove that ∀FO is superfluous with respect to the complexity classes **NL**, **P**, **NP** and **coNP**. We have enlarged that list proving that the superfluity method is also applicable in the complexity classes corresponding to the Second Level of the Polynomial-Time Hierarchy, Σ_2^p and Π_2^p.

As Definition 3 is strongly semantical, there is still no generic proof of superfluity in every level of **PH**, but we believe it is the case. One way to tackle this problem might be considering a sequence $\{A_k\}$ of Σ_k^p-complete problems, with an intrinsic relation on the vocabularies involved.

Immerman-Medina conjecture is still a source for future investigation, since no other fragments of FO has been proven superfluous. A superfluity version of ∃FO with respect to **NP** might lead to an inductive proof of superfluity for FO.

References

1. E. Allender, J. Balcazar and N. Immerman: *A first-order isomorphism theorem.* SIAM Journal of Computing, 26(2):555567, 1997.
2. N. Borges, B. Bonet: *Universal FO is superfluous for NL, P, NP and CO NP.* Logical Methods in Computer Science, Vol. 10(1:15), pp. 1-16. 2014.
3. N. Borges, B. Bonet: *On canonical forms of complete problems via first order projections.* LCC 07: Workshop on Logic and Computational Complexity. 2007.
4. J. Berit: *New Classes of Complete Problems for the Second Level of the Polynomial Hierarchy.* Doctoral Thesis. Technischen Universitat Berlin. 2011.
5. R. Fagin: *Generalized First-Order Spectra and Polynomial-Time Recognizable Sets.* Complexity of Computation, ed. R. Karp, SIAM-AMS Proc. 7, pp. 27-41. 1974.
6. D. Marx: *Complexity of clique coloring and related problems.* Theoret. Comput. Sci. 412(29), 34873500. 2011.
7. J. Medina: *A Descriptive Approach To The Class NP.* PhD thesis, University of Massachusetts, Amherst, 1997.
8. C. Papadimitriou: *Computational Complexity.* Addison-Wesley Publishing Company, 1st edition, 1995.
9. E. Pin, N. Borges: *A syntactic tool for proving hardness in the Second Level of the Polynomial-Time Hierarchy.* Preprint. arXiv:1707.09327. 2017.

A Problems in Σ_2^p

1. $2-$ Quantified Satisfiability (QSat_2)
   Instance: : A DNF Boolean formula $\phi(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ and $\mathbf{y}$ are tuples of Boolean variables of length n and m, respectively.
   Question: Is there a vector $\mathbf{x} \in \{0,1\}^n$ such that for every vector $\mathbf{y} \in \{0,1\}^m$, $\phi(\mathbf{x}, \mathbf{y})$ is true?
2. $2-$ Quantified Unsatisfiability (QUnsat_2)
   Instance: : A CNF Boolean formula $\phi(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ and $\mathbf{y}$ are tuples of Boolean variables of length n and m, respectively.
   Question: Is there a vector $\mathbf{x} \in \{0,1\}^n$ such that for every vector $\mathbf{y} \in \{0,1\}^m$, $\phi(\mathbf{x}, \mathbf{y})$ is false?
3. Unique Extension Satisfiability ($\exists\exists!\mathrm{Sat}$)
   Instance: : A CNF Boolean formula $\phi(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ and $\mathbf{y}$ are tuples of Boolean variables of length n and m, respectively.
   Question: Is there a vector $\mathbf{x} \in \{0,1\}^n$ such that for an unique vector $\mathbf{y} \in \{0,1\}^m$, $\phi(\mathbf{x}, \mathbf{y})$ is true?
4. $2-$ Clique Coloring (2CC)
   Instance: : A simple graph $\mathcal{G} = \langle V, E\rangle$.
   Question: Is there a 2-clique coloring of $\mathcal{G}$, that is, a map $c : V \to \{\text{red}, \text{blue}\}$ such that every maximal clique (complete subgraph) of $\mathcal{G}$ is nonmonochromatic?