Herbrand Property, Finite Quasi-Herbrand Models, and a Chandra-Merlin Theorem for Quantified Conjunctive Queries*

Fabio Mogavero
University of Oxford

A fundamental theorem by Skolem [24] establishes that every first-order sentence without equality is *satisfiable* if and only if its *functional (Skolem) normal form* has a *canonical (Herbrand) model*. In this context, the universe of discourse is the set of ground terms over the vocabulary of the sentence and the interpretation of the functions is defined in an algebraically transparent way: each term denotes precisely itself. A breakthrough technique by Büchi [9], further refined by Aanderaa [1] and Börger [5], exploits the structure of Herbrand models of relational first-order $\exists\forall\exists\forall$-sentences to prove the undecidability of the corresponding prefix class. Here, the Herbrand universe encodes the set of natural numbers with zero and successor. In that way, the data structure operated by a two-register machine are implemented transparently. This allows, therefore, an elementary reduction from the associated halting problem, which bypasses entirely the cumbersome axiomatization of the underlying register operations [6]. The transparency of classic Herbrand interpretations, which underlies their success as a tool for undecidability proofs, as well as numerous other applications in mathematical logic and theoretical computer science (e.g., in completeness theorems [15], semantic tableaux [19], alternative first-order semantics [16], automated reasoning [10], logic programming [21], and database theory [2]), comes at a price: their lack of succinctness. Indeed, as soon as the vocabulary contains a function symbol, the corresponding Herbrand universe becomes infinite. This phenomenon severely limits the effectiveness of Herbrand models in establishing the decidability of fragments of first-order logic with functions, not to mention in obtaining tight computational-complexity bounds or model-theoretic results like the finite-model property. Aiming at decidability, however, more useful appears a property of Herbrand models implied by their transparency, rather than the transparency itself: *an equality between terms is satisfiable on a Herbrand model if and only if its terms are unifiable*. Intuitively, the particular interpretation of terms neutralizes the expressive power of equalities, by reducing their satisfiability, at first glance a hard, even infinitary, question, to a polynomial-time unifiability test. This observation has been exploited by Kozen to show that the validity problem of positive first-order logic is in NPTIME [20].

The present work is devoted to the study and application of the *Herbrand property*, a novel model-theoretic notion expressing the fact that the satisfiability of an equation boils down to the unifiability of its terms. In this terminology, the aforementioned observation by Kozen can be rephrased as follows: every Herbrand model enjoys the Herbrand property. Our work, though, tackles the concept per se, abstracting it from the specific implementation via Herbrand models, and investigates its consequences from both a finite model-theoretic and a structural-complexity perspective. We obtain non-trivial results on both levels.

A first part of the work is devoted to the finite model-theoretic development of the Herbrand property. The main result is a universal and finitary version of this concept, as follows.

> *A set of terms is equalizable over all finite structures if and only if it is unifiable.*

Here, a set of terms is said to be *equalizable* over a structure if such a structure satisfies all pairwise equalities between terms of the set. The universal and finitary aspect of our characterization contrasts with the reduction observed by Kozen from *equalizability over Herbrand models* to unifiability, as we reduce *equalizability on all finite structures* to unifiability. An easy corollary of the above result is the existence of finite models enjoying the Herbrand property, which we call *finite quasi-Herbrand models*. This can be seen as an evidence of the fact that the intrinsic infinitary nature of Herbrand models over vocabularies with functions is inessential. In other words, the latter can be seen as a naively verbose implementation of this fundamental concept. The main consequence of our finitary version of the Herbrand property is that satisfiable universal single-binding sentences have finite quasi-Herbrand models, i.e., more abstractly, the fragment of universal single-binding logic enjoys the *finite* (technically, *small* [25]) *model property*.

Universal single-binding logic is the language of positive Boolean combinations of universally quantified binding forms, where a *binding form* is, in turn, a Boolean combination of relational atoms over the same tuple of terms. This logic is syntactically contained in *conjunctive-binding logic* introduced in [22], a fragment of first-order logic that allows positive Boolean combinations of quantified conjunctions of binding forms. Since the satisfiability problems for the two logics are succinctly interreducible via skolemization, we have

> *Conjunctive-binding logic enjoys the finite model property.*

In particular, its (finite) satisfiability problem is decidable, answering an open question in the literature and completing the decidability classification of binding fragments of first-order logic [22]. The result can also be read as a non-trivial generalization of the decidability proof for *Herbrand logic* [14], the language of quantified conjunctions of literals, as it is syntactically contained in the logic under analysis. On the other hand, conjunctive-binding logic is orthogonal to all known decidable fragments (prefix classes [6], two variable [18], [23], guarded fragments [3], [17], guarded negation [4], et cetera, see [22] for details) and its solution requires different ideas and techniques.

The rest of the work focuses on the consequences of the Herbrand property from the structural complexity viewpoint with respect to various satisfiability and entailment problems in conjunctive-binding logic and fragments thereof. Our first

* Joint work with Simone Bova, Vienna University of Technology, to be published in the proceedings of LICS 2017.

result is a characterization of the (finite) satisfiability problem for universal single-binding logic in terms of (finite) quasi-Herbrand models, placing this problem at the third level of the polynomial hierarchy. The aforementioned interreducibility allows then to prove the following statement.

The (finite) satisfiability problem for conjunctive-binding logic is Σ_3^{P}-complete.

As opposed to satisfiability, the entailment problem for conjunctive-binding logic is, unfortunately, undecidable. Interestingly enough, the prominent syntactic fragment of *quantified conjunctive queries (*QCQ*)* has been shown to have a decidable (*general*) entailment problem by Chen, Madelaine, and Martin [12]. This problem is closely related to QCQ *containment* in database theory. In this context, however, the notion of interest is *finite* entailment, i.e., entailment on all finite structures, as in most applications the database is finite. The question whether entailment and finite entailment in QCQ coincide, though, was left open in [12]. Our second result is a tight structural complexity classification of general and finite entailment in *positive Herbrand logic (*PH*)*, the logic of quantified conjunctions of *atoms*, which syntactically contains QCQ.

The (finite) entailment problem in positive Herbrand logic is NPTIME*-complete.*

Our result has both a complexity-theoretic and a logical value: on the one hand it closes the previously standing gap between NPTIME-hardness and 3EXPTIME-membership for QCQ containment [12]; on the other hand, by exploiting our *finitary Herbrand property*, it actually pushes the finite version of the problem in NPTIME, even for PH. In retrospect, and not coincidentally, Chen, Madelaine, and Martin obtain their 3EXPTIME upper bound by reasoning on a finite substructure of an infinite Herbrand model associated with the Skolem normal form of the implicant sentence in the instance. Our proof of this theorem, placing the problem in NPTIME, relies on the observation that positive instances of PH entailment have short resolution refutations. A careful inspection reveals that such small witnesses encode certain mappings from the consequent to the antecedent in the instance. In particular, in the special case of conjunctive queries (CQ), it is readily seen that these mappings are precisely homomorphisms. We have thus recovered the classic theorem by Chandra and Merlin [11], which places the (finite) containment question for CQ in NPTIME. Our third and final result, stemmed from this insight, consists in abstracting a lifted notion of homomorphism from short refutations of positive QCQ entailment instances. This notion characterizes the QCQ containment problem.

Given two QCQ*s ϕ and ψ, it holds that $\phi \models \psi$ if and only if ψ admits a Skolem homomorphism to ϕ.*

A *Skolem homomorphism* is a substitution of the variables in ψ by terms of the skolemization of ϕ, which is both sensitive to the dependencies induced by the quantifier prefix of ψ and faithful to the relational structure associated with ϕ. Besides, such an alleged Skolem homomorphism is efficiently checkable relative to ϕ and ψ, thus yielding an alternative view on the NPTIME-membership of the QCQ containment problem. Our result can be read, therefore, as an accurate lifting of Chandra-Merlin theorem to the QCQ realm.

Placing the (finite) entailment problem for PH within NPTIME not just closes the wide complexity-theoretic gap between the NPTIME-hardness of CQ containment [11] and the 3EXPTIME-membership of QCQ containment [12], it actually pushes the problem in the range of practically feasible computation, e.g., via SAT solvers. Interestingly, resolution-based first-order provers, once executed on QCQ-containment instances, implement in essence the behavior dictated by the proposed extension of the Chandra-Merlin theorem.

We believe that the ideas in this work have the potential for nontrivial developments. An intriguing problem is the rewriting of QCQs in order to minimize the number of distinct variables used in the query, which is the known algorithmic bottleneck for query evaluation. The issue, fully understood on CQs [7], [8], [13], is wide open on QCQs. Indeed, the problem is not even known to be decidable. Perhaps, the notion of Skolem homomorphism might eventually offer a viable approach.

REFERENCES

[1] S. Aanderaa, “On the Decision Problem for Formulas in which all Disjunctions are Binary.” in *SLS’71*. Elsevier, 1971, vol. 63, pp. 1–18.
[2] S. Abiteboul, R. Hull, and V. Vianu, *Foundations of Databases.* Addison-Wesley, 1995.
[3] H. Andréka, J. van Benthem, and I. Németi, “Modal Languages And Bounded Fragments Of Predicate Logic.” *JPL*, vol. 27, no. 3, pp. 217–274, 1998.
[4] V. Bárány, B. ten Cate, and L. Segoufin, “Guarded Negation.” *JACM*, vol. 62, no. 3, pp. 22:1–26, 2015.
[5] E. Börger, “Reduktionstypen in Krom- und Hornformeln.” Ph.D. dissertation, University of Münster, Münster, Germany, 1971.
[6] E. Börger, E. Grädel, and Y. Gurevich, *The Classical Decision Problem.*, ser. Perspectives in Mathematical Logic. Springer, 1997.
[7] S. Bova and H. Chen, “The Complexity of Width Minimization for Existential Positive Queries.” in *ICDT’14*. OpenProceedings.org, 2014, pp. 235–244.
[8] ——, “How Many Variables are Needed to Express an Existential Positive Query?” in *ICDT’17*. OpenProceedings.org, 2017, pp. 9:1–16.
[9] J. Büchi, “Turing-Machines and the Entscheidungsproblem.” *MA*, vol. 148, no. 3, pp. 201–213, 1962.
[10] C.-L. Chang and R.C.T. Lee, *Symbolic Logic and Mechanical Theorem Proving.*, ser. Computer Science Classics. Academic Press, 1973.
[11] A. Chandra and P. Merlin, “Optimal Implementation of Conjunctive Queries in Relational Data Bases.” in *STOC’77*. ACM, 1977, pp. 77–90.
[12] H. Chen, F. Madelaine, and B. Martin, “Quantified Constraints and Containment Problems.” *LMCS*, vol. 11, no. 3, pp. 1–28, 2015.
[13] V. Dalmau, P. Kolaitis, and M. Vardi, “Constraint Satisfaction, Bounded Treewidth, and Infinite-Variable Logics.” in *CP’02*, ser. LNCS 2470. Springer, 2002, pp. 310–326.
[14] L. Denenberg and H. Lewis, “The Complexity of the Satisfiability Problem for Krom Formulas.” *TCS*, vol. 30, no. 3, pp. 319–341, 1984.
[15] H.-D. Ebbinghaus, J. Flum, and W. Thomas, *Mathematical Logic.*, ser. Undergraduate Texts in Mathematics. Springer, 1984.
[16] M. Genesereth and E. Kao, *Introduction to Logic.*, ser. Synthesis Lectures on Computer Science. Morgan & Claypool Publishers, 2013.
[17] E. Grädel, “Decision Procedures for Guarded Logics.” in *CADE’99*, ser. LNCS 1632. Springer, 1999, pp. 31–51.
[18] E. Grädel, P. Kolaitis, and M. Vardi, “On the Decision Problem for Two-Variable First-Order Logic.” *BSL*, vol. 3, no. 1, pp. 53–69, 1997.
[19] R. Jeffrey, *Formal Logic: Its Scope and Limits.* McGraw-Hill, 1981.
[20] D. Kozen, “Positive First-order Logic is NP-complete.” *IBMJRD*, vol. 25, no. 4, pp. 327–332, 1981.
[21] J. Lassez, M. Maher, and K. Marriott, “Unification Revisited.” in *FLFP’86*, ser. LNCS 306. Springer, 1988, pp. 67–113.
[22] F. Mogavero and G. Perelli, “Binding Forms in First-Order Logic.” in *CSL’15*, ser. LIPIcs 41. Leibniz-Zentrum fuer Informatik, 2015, pp. 648–665.
[23] M. Mortimer, “On Languages with Two Variables.” *MLQ*, vol. 21, no. 1, pp. 135–140, 1975.
[24] T. Skolem, “Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen.” vol. 4, 1920.
[25] A. Urquhart, “Decidability and the Finite Model Property.” *JPL*, vol. 10, no. 3, pp. 367–370, 1981.