Annals of Pure and Applied Logic 164 (2013) 907–927

Contents lists available at SciVerse ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

Super/rosy L^k-theories and classes of finite structures

Cameron Donnay Hill *

University of Notre Dame, Department of Mathematics, 255 Hurley, Notre Dame, IN 46556, United States

ARTICLE INFO

Article history:
Received 25 July 2011
Received in revised form 22 January 2013
Accepted 10 April 2013
Available online 5 June 2013

MSC:
03B70
03C13
03C15
03C45

Keywords:
Rosy theory
Finite-variable logic
Finite structures
Amalgamation
þ-Independence
þ-Forking

ABSTRACT

We recover the essentials of þ-forking, rosiness and super-rosiness for certain amalgamation classes K, and thence of finite-variable theories of finite structures. This provides a foundation for a model-theoretic analysis of a natural extension of the "L^k-Canonization Problem" – the possibility of efficiently recovering finite models of T given a finite presentation of an L^k-theory T. Some of this work is accomplished through different sorts of "transfer" theorem (of varying degrees of subtlety) to the first-order theory $T^{\lim}$ of the direct limit. Our results include, to start with, a recovery of the basic technology of þ-independence (analogous to Onshuus (2006) [15]) using a rather straightforward transfer. We also recover an analog of the "þ-Independence theorems" of Ealy and Onshuus (2007) [7] for amalgamation classes and their limits by showing how to transfer/lift an abstract independence relation $\mathop{\smile\!\!\!\!\!\!\mid}^{\circ}$ on the amalgamation class to the limit theory $T^{\lim}$. We also work out an appropriate notion of Local Character for independence relations over classes finite structures, and we use this to verify that rosiness and super-rosiness-with-finite-$U^{þ}$-ranks coincide in these amalgamation classes and their limit theories.

Introduction

This article is the first of a three-part series (with [10] and [11]) examining the intrinsic geometry of an algorithmic problem – the L^k-Canonization Problem – which is well-known to finite-model theorists and some complexity theorists. L^k denotes the fragment of first-order logic consisting of formulas with at most k distinct variables, *free or bound*, and it can be shown that for any finite structure $\mathcal{M}$ (in a finite relational signature), its complete k-variable theory $Th^k(\mathcal{M})$ is finitely axiomatizable in a uniform way. The L^k-Canonization Problem asks us to devise an operator F that takes the theories $Th^k(\mathcal{M})$ to finite models $F(Th^k(\mathcal{M})) \models Th^k(\mathcal{M})$ – thus, defining a "canonical" model of each complete k-variable theory that does have finite models. Composing the canonization operator F with the mapping $\mathcal{M} \mapsto Th^k(\mathcal{M})$, the operator $F(Th^k(-))$ can be thought of as a solution to a natural relaxation of the Graph Isomorphism Problem, the status of which is a major open problem in complexity theory (see [19] for an old survey).

Although this problem is certainly unsolvable over the class of *all* L^3-theories [8], it has been shown that for the class of stable L^k-theories and for the class of super-simple L^k-theories with trivial forking dependence (with additional amalgamation assumptions), the L^k-Canonization Problem is *recursively* solvable (see [4] and [5], respectively). In both of those cases, resolution of the L^k-Canonization Problem is reduced to showing that certain complete first-order theories associated with the original L^k-theories have the finite sub-model property. Thus, after the heavy lifting done by the model theory, the algorithm itself is extremely simple-minded. Moreover, the analyses in [4] and [5] do not assume *a priori* that the L^k-theories in

* Tel.: +1 574 631 7776.
E-mail address: cameron.hill.136@nd.edu.

0168-0072/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.apal.2013.04.003

question certainly have finite models. In contrast, in this series of articles, we will examine the L^k-Canonization Problem for L^k-theories that do certainly have finite models. Moreover, we will consider implementation of L^k-Canonization operators in a significantly restricted model of computation, leading to a notion we call "efficient constructibility." Finally, we will expand the original L^k-Canonization Problem to take whole L^k-elementary diagrams as input, which allows us to work with individual L^k-theories in a non-trivial way. Thus, the goal of this series of articles is to prove the following:

Main Theorem. *Let $\mathcal{M}_0$ be a finite structure, and let K be the class of all finite models of $T = Th^k(\mathcal{M}_0)$. Assuming that K has adequate amalgamation properties, let $T^{\lim}$ be the complete first-order theory of the direct limit of K. Then the following are equivalent:*

1. *$T^{\lim}$ is super-rosy of finite $U^{þ}$-rank.*
2. *K is rosy.*
3. *K is efficiently constructible – meaning that in a certain weak model of computation (exposed in [10]), the following problem is computable*:

> **K**-***Construction*** *problem*:
> Given $\mathcal{M}[A]$ *for some (implicit)* $\mathcal{M} \in K$ *and* $A \subseteq M$,[1]
> Return $\mathcal{N} \in K$ *such that* $A \subseteq N$ *and* $\mathcal{N}[A] = \mathcal{M}[A]$.

The equivalence of (1) and (2) in the main theorem is, indeed, intuitively obvious, and this suggests that the statement, "K is rosy" may have no substantial content of its own. However, proving the equivalence with (3) seems to require that we first make sense of rosiness and þ-independence in K in its own right. In particular, the proof requires a characterization of rosiness by abstract independence relations that accommodate *only* triples of finite sets. Thus, in this article, we will settle on what is meant by "K is rosy" and work out just how the rosiness of K and that of $T^{\lim}$ interact. In a final section of the article, we will also see that much of the development does not really require the context of k-variable logic – the more general context of "super-robust classes" of finite structures is actually sufficient.

1. Background and the main setting

1.1. Finite-variable logics

Finite-variable fragments of first-order logic, L^k, were formulated by many authors independently (e.g. [18], but our main references have been [16] and [13]). The importance of L^k and its infinitary extension $L^k_{\infty,\omega}$ in finite-model theory is difficult to overstate. For our purposes, L^k is satisfying because a "complete" L^k-theory – that is, complete for L^k-sentences – can have many non-isomorphic finite models, which is surely a prerequisite for bringing classical model-theoretic ideas to bear in finite-model theory.

Definition 1.1. Let ϱ be a finite relational signature. Assume $k \geqslant \mathrm{ari}(\varrho) = \max\{\mathrm{ari}(R)\colon R \in \varrho\}$ and $k \geqslant 2$.

1. Fix a set $X = \{x_1, \ldots, x_k\}$ of exactly k distinct variables. Then, $L^X = L^X_\varrho$ is the fragment of the first-order logic $L = L_\varrho$ keeping only those formulas all of whose variables, *free or bound*, come from X. If $V = \{x_0, x_1, \ldots, x_n, \ldots\}$ is the infinite set of first-order variables understood in the construction of the full first-order logic, then $L^k = \bigcup\{L^X\colon X \in \binom{V}{k}\}$, where $\binom{V}{k}$ is the set of k-element subsets of V.
   As usual, we write $\phi(x_1, \ldots, x_k)$ to mean that the set of free variables of ϕ is a subset of $\{x_1, \ldots, x_k\}$, but not necessarily identical to it.
2. For a ϱ-structure $\mathcal{M}$, the k-variable theory of $\mathcal{M}$, denoted $Th^k(\mathcal{M})$ is the set of sentences ϕ of L^k such that $\mathcal{M} \models \phi$. Note that $Th^k(\mathcal{M})$ is complete with respect to k-variable sentences in that either $\phi \in Th^k(\mathcal{M})$ or $\neg\phi \in Th^k(\mathcal{M})$ for every k-variable sentence ϕ.
3. For a k-tuple $\bar{a} \in M^k$, we set $tp^k(\bar{a}; \mathcal{M}) = \{\phi(x_1, \ldots, x_k) \in L^k\colon\ \mathcal{M} \models \phi(\bar{a})\}$ and if $T = Th^k(\mathcal{M})$, then $S^k_k(T) = \{tp^k(\bar{a}; \mathcal{M})\colon \bar{a} \in M^k\}$.
   It can be shown – in a number of ways – that for a complete L^k-theory T, T has a finite model only if $S^k_k(T)$ is finite. All of those methods also show that $S^k_k(T)$ is an invariant of $T = Th^k(\mathcal{M})$ rather than $\mathcal{M}$ itself – that is, if $Th^k(\mathcal{N}) = T$ for some other ϱ-structure $\mathcal{N}$ (equivalently, if $\mathcal{N} \equiv^k \mathcal{M}$), then $\{tp^k(\bar{b}; \mathcal{N})\colon \bar{b} \in N^k\} = S^k_k(T)$, too. Finally, it can also be shown that if $\mathcal{M}$ is finite, then $Th^k(\mathcal{M})$ is finitely axiomatizable, and in fact, the mapping $\mathcal{M} \mapsto Th^k(\mathcal{M})$ is computable in Rel-Ptime (see [1]). This latter fact is the basis for our notion of efficient constructibility.
4. Let $\mathcal{M}$ be a ϱ-structure, and let $B \subseteq M$. Then for $X \in \binom{V}{k}$ as above and $e : X \to B \cup X$, we define $L^X(e)$ to be the set,

$$\{\phi(e(x_1), \ldots, e(x_n))\colon \phi(x_1, \ldots, x_n) \in L^X\}.$$

[1] Here $\mathcal{M}[A]$ denotes the induced substructure of $\mathcal{M}$ on the subset A, so $\mathcal{M}[A]$ *does not* carry any further information about $\mathcal{M}$.