Two-Variable Logic with Counting is Decidable*

Erich Grädel Martin Otto Eric Rosen
Mathematische Grundlagen der Informatik
RWTH Aachen, Germany
{graedel,otto,erosen}@informatik.rwth-aachen.de

Abstract

We prove that the satisfiability and the finite satisfiability problems for C^2 are decidable.

C^2 is first-order logic with only two variables in the presence of arbitrary counting quantifiers $\exists^{\geqslant m}$, $m \geqslant 1$. It considerably extends L^2, plain first-order with only two variables, which is known to be decidable by a result of Mortimer's. Unlike L^2, C^2 does not have the finite model property.

1. Introduction

Let L^2 be the fragment of first-order logic (with equality), that only has the variable symbols x and y — i.e., the closure of atomic formulae involving no variables apart from x and y under Boolean operations and $\exists x$, $\exists y$. Throughout this paper we restrict attention to finite vocabularies consisting of relation symbols and constants. By [8], L^2 has the *finite model property*, i.e., every satisfiable sentence has a finite model. Consequently, the satisfiability problem for L^2 is decidable.

Standard terminology uses the following:

- *sat*(X) for the set of $\psi \in X$ that are satisfiable;
- *fin-sat*(X) for the set of $\psi \in X$ that have a finite model;
- *inf-sat*(X) for the set of $\psi \in X$ that have an infinite model;
- *inf-axioms*(X) for *sat*(X) − *fin-sat*(X), the *infinity axioms* of X.

By the finite model property, *sat*(L^2) = *fin-sat*(L^2). Note that *fin-sat*(X) is recursively enumerable for any formula class with effective semantics, and that *sat*(X) is co-r.e. for any X that can effectively be embedded into first-order logic FO (by completeness). Thus, as a fragment of FO with the finite model property, L^2 is decidable. In fact, from Mortimer's proof it follows that every satisfiable L^2-sentence has a model of double exponential size, so there is a more direct argument for decidability. Mortimer's proof and the quality of this bound were recently improved upon in [4], where the complexity of *sat*(L^2) was also determined precisely: *sat*(L^2) is complete for nondeterministic exponential time.

Some of the important applications of this result arise in the context of modal logics that can be embedded into L^2. From a practical point of view, L^2 remains too weak for many applications, though. Although several extensions of modal logic like propositional dynamic logic PDL, computation tree logic CTL, or propositional μ-calculus L_μ (which feature expressive means that make them useful as process logics) are known to be decidable, it was shown in [5] that several corresponding extensions of L^2 are no longer decidable.

Unlike FO, L^2 is not closed under assertions that there are at least m elements satisfying some property for $m > 2$. It is therefore natural to extend L^2 to allow arbitrary counting quantifiers $\exists^{\geqslant m}$, $m \geqslant 1$.

Definition 1.1. *C^2 is the extension of L^2 which admits all counting quantifiers $\exists^{\geqslant m}$, $m \geqslant 1$, rather than just $\exists$.*

For instance, the C^2-sentence $\forall x \exists^{\leqslant m} y Exy$ defines the class of all graphs whose degree is bounded by m. Here we have already allowed derived quantifiers $\forall$ for $\neg\exists^{\geqslant 1}\neg$ and $\exists^{\leqslant m}$ for $\neg\exists^{\geqslant m+1}$. Quantifiers $\exists^{=m}$, $\exists^{>m}$ etc. may similarly be admitted in C^2 without increasing its expressive power.

We point out one particularly interesting piece of evidence for the expressive power of C^2, which far exceeds that of L^2. Immerman and Lander [7] show that the C^2-theory of a finite graph (which is actually axiomatized by a single sentence of C^2) exactly charac-

* This work has been partially supported by the German-Israeli Foundation of Scientific Research and Development.

terises the stable colouring of that graph. By a result of Babai, Erdös, and Selkow [1] it follows that almost all finite graphs are characterized up to isomorphism by their C^2-theory — almost all in the sense of asymptotic probabilities: the proportion of graphs with vertices $0, \ldots, n-1$ having this property tends to 1 as n goes to infinity. For more information on the expressive power of and model-theoretic properties of C^2 (and extensions thereof) we refer to [9, 10].

That C^2 does not have the finite model property is witnessed by

$$\forall x \exists^{=1} y Exy \wedge \forall x \exists^{\leqslant 1} y Eyx \wedge \exists x \forall y \neg Eyx,$$

which asserts that E is the graph of an injective function from the universe to itself which fails to be surjective. This is clearly an infinity axiom. Therefore *fin-sat*$(C^2) \subsetneq$ *sat*(C^2), and there are two different decidability issues to be settled.

Since counting quantifiers can be simulated in first-order logic (albeit with an increase in the number of variables) C^2 is a natural fragment of first-order logic. Thus both *sat*(C^2) and *fin-sat*(C^2) are instances of the *classical decision problem* of mathematical logic, i.e., the problem of classifying fragments of first-order logic according to whether their satisfiability and finite satisfiability problems are decidable (see [3]).

These are the main results of the paper:

Theorem 1.2. *sat*(C^2) *is decidable.*

Theorem 1.3. *fin-sat*(C^2) *is decidable.*

In the present exposition we give a full proof for the decidability of *sat*(C^2). As $C^2 \subseteq$ FO, it suffices to establish that *sat*(C^2) is r.e., and for this — as *fin-sat*(C^2) clearly is r.e. — it suffices to show that *inf-sat*(C^2) is r.e. We also provide a detailed sketch of the proof for the decidability of *fin-sat*(C^2). The complete proof will appear in the full version of the paper. Further, there we shall also point out some consequences of our decidability results, e.g., for description logics with counting restrictions and modal logics with graded modalities.

Remark. An upper complexity bound for *sat*(C^2) has recently been announced by Pacholski, Szwast and Tendera [11].

2. Notation and basic definitions

Let τ be some finite vocabulary of unary and binary relation symbols and constants. It is not difficult to see that in two-variable logics (even in the presence of constants) predicates of higher arity can be eliminated with only a linear increase in formula length (see [4]).

We shall often suppress τ in our notation. If K is any set of parameters, we may treat these as new constant symbols and write $\tau_K = \tau \dot{\cup} K$ for the extended vocabulary. Usually K is a subset of some structure, and we do not distinguish between the element $k \in K$ and the constant k which is the syntactic name for k. We denote structures as $\mathfrak{A} = (A, \ldots)$ where A is the universe of $\mathfrak{A}$.

An *atomic n-type* over vocabulary τ is a maximally consistent set of atomic and negated atomic τ-formulae in n variable symbols.

Let $\boldsymbol{\alpha}$ denote the finite set of atomic 1-types over τ in the single variable x. Let $\boldsymbol{\beta}$ stand for the finite set of those atomic 2-types over τ in two variables x and y which include $\neg x = y$. $\boldsymbol{\alpha}^K$ will denote the set of atomic 1-types over τ_K in the variable x. $\boldsymbol{\alpha}^K$ is finite if K is.

We use letters α and β to denote typical elements of $\boldsymbol{\alpha}$ (or $\boldsymbol{\alpha}^K$) and $\boldsymbol{\beta}$, respectively. For a τ-structure $\mathfrak{A}$, some fixed subset $K \subseteq A$, and distinct elements a and b in A let

$$\begin{array}{lcl} \mathrm{atp}_{\mathfrak{A}}(a) & \in & \boldsymbol{\alpha}, \\ \mathrm{atp}^K_{\mathfrak{A}}(a) & \in & \boldsymbol{\alpha}^K, \\ \mathrm{atp}_{\mathfrak{A}}(a,b) & \in & \boldsymbol{\beta} \end{array}$$

denote the respective atomic types. It is clear that for $\alpha \in \boldsymbol{\alpha}$ and quantifier-free η in the single variable x, it can effectively be determined whether $\alpha \models \eta$, i.e., whether realizations of α necessarily satisfy η. (To see this put η in disjunctive normal form.) Similarly $\beta \models \eta$, for $\beta \in \boldsymbol{\beta}$ and η quantifier-free, and $\alpha \models \eta$, for $\alpha \in \boldsymbol{\alpha}^K$ and quantifier-free η in the single variable x but possibly with parameters from K, are decidable.

We shall see that the analysis of structures with respect to C^2 can make good use of one further fundamental notion of type, which explicitly incorporates some limited counting information.

Definition 2.1. *For an element a of $\mathfrak{A}$ let the* counting star *of a, denoted* $\mathrm{ctp}_{\mathfrak{A}}(a)$, *be the function*

$$\begin{array}{rcl} \gamma\colon \boldsymbol{\beta} & \longrightarrow & \{0,1,2^+\} \\ \beta & \longmapsto & \dot{\#}_{b \in A}\big(\mathrm{atp}_{\mathfrak{A}}(a,b) = \beta\big). \end{array}$$

$\dot{\#}_{x \in S} \ldots$ counts the number of $x \in S$ satisfying $\ldots$ according to 0, 1, 2^+ (= many). We use $\boldsymbol{\gamma}$ to denote the finite set of all non-degenerate satisfiable counting stars, i.e., of all $\mathrm{ctp}_{\mathfrak{A}}(a)$ that are realized in some structure $\mathfrak{A}$ that has at least two elements. For any particular $\mathfrak{A}$ (with at least two elements) let $\boldsymbol{\gamma}_{\mathfrak{A}} \subseteq \boldsymbol{\gamma}$ denote the set of those $\gamma \in \boldsymbol{\gamma}$ that are realized in $\mathfrak{A}$.

Any atomic 2-type $\beta \in \boldsymbol{\beta}$ uniquely determines two atomic 1-types $\mathrm{atp}_1(\beta)$ and $\mathrm{atp}_2(\beta)$ as the 1-types of

x and y respectively that are prescribed in β. Note that via $\mathrm{atp}_1(\beta)$, for those β with $\gamma(\beta) \neq 0$, each $\gamma \in \boldsymbol{\gamma}$ determines *its atomic 1-type* $\mathrm{atp}(\gamma)$. Semantically $\mathrm{atp}(\gamma)$ is the unique $\alpha \in \boldsymbol{\alpha}$ that is realized by all a that realize γ.

We shall also make use of the mapping on atomic 2-types β that exchanges the roles of x and y:

$$\beta \longmapsto \overline{\beta} := \beta\tfrac{yx}{xy}.$$

The truncated counting information in the counting stars plays an essential role with respect to the following useful normal form, which only involves counting up to 2.

2.1. A normal form

Definition 2.2. *Say that a sentence $\varphi \in C^2$ is in normal form if it is a conjunction of sentences of the following kinds: $\forall x \forall y \eta$ and $\forall x \exists^{=1} y \eta$, where the η are quantifier-free.*

Lemma 2.3. *There is a recursive reduction* NF *from C^2-sentences to C^2-sentences in normal form (over an extended vocabulary), which is sound for satisfiability: $\varphi \in sat(C^2)$ if and only if $\mathrm{NF}(\varphi) \in sat(C^2)$. Likewise, it is also sound for finite satisfiability.*

Proof. The proof is given in two parts. We first show that for any sentence $\varphi \in C^2$ of vocabulary τ (which without loss of generality contains a constant c) there are sentences φ_0 and θ in an expanded vocabulary satisfying the following.

(i) θ is a conjunction of sentences of the form $\forall x \exists^{\geqslant m} y \eta$ and $\forall x \exists^{<m} y \eta$ for quantifier-free η and $m \geqslant 1$. φ_0 is quantifier-free.

(ii) Each τ-structure has a unique expansion to a model of θ.

(iii) θ implies equivalence of φ and φ_0.

Note that (ii) and (iii) imply that φ is satisfiable if and only $\theta \wedge \varphi_0$ is satisfiable. In the second step we shall transform $\theta \wedge \varphi_0$ into normal form to finish the proof of the lemma.

θ and φ_0 are constructed inductively with respect to the number of quantifiers in φ. If φ is quantifier-free, we are done. Otherwise consider a subformula ψ of type $\exists^{\geqslant m} y \chi$, where y is free in χ, but x may or may not be free in χ. These two cases are treated separately. Consider firstly $\psi(x) = \exists^{\geqslant m} y \chi(x, y)$, with displayed variables occurring free. Introduce a new unary predicate P and let θ_1 be the conjunction of

$$\forall x \exists^{\geqslant m} y \big(Px \rightarrow \chi(x, y)\big)$$
$$\forall x \exists^{<m} y \big(\neg Px \wedge \chi(x, y)\big)$$

which is equivalent to $\forall x \big(Px \leftrightarrow \exists^{\geqslant m} y \chi(x, y)\big)$. Let φ' be the result of replacing the subformula $\psi(x)$ in φ by the atom Px. Then φ' has fewer quantifiers than φ, θ_1 is of the desired form, and, since any model of θ_1 must interpret P as $\{x | \psi(x)\}$, it follows that $\theta_1 \models \varphi \leftrightarrow \varphi'$.

If $\psi = \exists^{\geqslant m} y \chi(y)$ does not have x as a free variable, then one may similarly use a unary P and the constant c to simulate a Boolean value in quantifier-free fashion. We may take the conjunction of the following for θ_1:

$$\forall x \exists^{\geqslant m} y \big(Px \rightarrow \chi(y)\big)$$
$$\forall x \exists^{<m} y \big(\neg Px \wedge \chi(y)\big).$$

θ_1 forces P to be the full, respectively empty, predicate according to the truth value of ψ. For φ' we now take the result of substituting the atom Pc for ψ in φ. Again $\theta_1 \models \varphi \leftrightarrow \varphi'$.

An inductive application of this procedure eventually yields θ (as the conjunction of the θ_i of each step) and a quantifier-free φ_0, as desired.

It remains to transform a sentence $\theta \wedge \varphi_0$ as obtained into proper normal form without affecting satisfiability. As φ_0 cannot have any free variables, it may as well be universally quantified to form an $\forall\forall$-conjunct. By using $\forall\forall$-conjuncts to eliminate other quantifier-free constituents, we may actually assume that the quantifier-free parts of the $\forall\exists^{\geqslant m}$- and $\forall\exists^{<m}$-conjuncts in θ are atomic formulae Pxy.

In order to translate $\forall x \exists^{\geqslant m} y Pxy$ into normal form we use m new binary predicates $P_1, \ldots, P_m$ and the conjunction of

$$\forall x \forall y \big(\bigvee_i P_i xy \ \rightarrow Pxy\big)$$
$$\bigwedge_{i \neq j} \forall x \forall y \big(P_i xy \rightarrow \neg P_j xy\big)$$
$$\bigwedge_i \forall x \exists^{=1} y \, P_i xy.$$

For $\forall x \exists^{<m} y Pxy$, where $m > 1$ we similarly use new binary $P_1, \ldots, P_{m-1}$ and the sentence

$$\forall x \forall y \big(Pxy \rightarrow \bigvee_i P_i xy\big) \wedge \bigwedge_i \forall x \exists^{=1} y \, P_i xy.$$

$\forall x \exists^{<1} y Pxy$, finally is equivalent to $\forall x \forall y \neg Pxy$ which is in normal form.

Clearly these replacements are sound for satisfiability and yield a sentence in normal form. ⊣

By the preceding lemma, we can restrict attention to sentences in normal form. For instance, for the proof of Theorem 1.2, it suffices to show that the set of sentences in *inf-sat*(C^2) in normal form is r.e.

This reduction is particularly important, since counting stars alone determine which sentences in normal form are satisfied in a structure.

Lemma 2.4. *Given φ in normal form and $\boldsymbol{\Upsilon}_{\mathfrak{A}}$ (for any $\mathfrak{A}$ with at least two elements), it is effectively decidable whether $\mathfrak{A} \models \varphi$.*

Proof. Recall that for $\beta \in \boldsymbol{\beta}$ and quantifier-free $\eta(x,y)$ it can directly be checked whether $\beta \models \eta(x,y)$ and whether $\mathrm{atp}_1(\beta) \models \eta(x,x)$. And similarly, for $\gamma \in \boldsymbol{\Upsilon}$ it can be checked whether $\mathrm{atp}(\gamma) \models \eta(x,x)$.

Consider now separately the constituent sentences of φ in normal form. For an $\forall\forall$-sentence $\forall x \forall y \eta(x,y)$ it suffices to check that for all $\gamma \in \boldsymbol{\Upsilon}_{\mathfrak{A}}$ and any β for which $\gamma(\beta) > 0$ it is true that $\beta \models \eta(x,y)$ and $\mathrm{atp}_1(\beta) \models \eta(x,x)$.

For an $\forall\exists^{=1}$-sentence $\forall x \exists^{=1} y \eta(x,y)$ similarly it merely has to be checked that all $\gamma \in \boldsymbol{\Upsilon}_{\mathfrak{A}}$ satisfy the following: either $\sum_{\beta \models \eta(x,y)} \gamma(\beta) = 1$ and $\mathrm{atp}(\gamma) \models \neg\eta(x,x)$, or $\mathrm{atp}(\gamma) \models \eta(x,x)$ and $\sum_{\beta \models \eta(x,y)} \gamma(\beta) = 0$. For the correctness of these sums note that different β are mutually exclusive. ⊣

3. Analysis of infinite structures

We turn to the analysis of infinite structures $\mathfrak{A}$. In a definable way we shall separate $\mathfrak{A}$ into a finite part and a rather homogeneous infinite part. We shall find *finite descriptions* for the infinite part that suffice to check for satisfaction of sentences in normal form.

Let $\mathfrak{A}$ be a fixed infinite τ-structure. Recall that $\boldsymbol{\Upsilon}$ is the set of all counting stars. Split $\boldsymbol{\Upsilon}$ into three disjoint subsets

$$\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}_0 \,\dot\cup\, \boldsymbol{\Upsilon}_{\mathrm{fin}} \,\dot\cup\, \boldsymbol{\Upsilon}_{\mathrm{inf}}$$

according to whether γ is realized in $\mathfrak{A}$ not at all, or finitely often, or infinitely often. Thus, the set $\boldsymbol{\Upsilon}_{\mathfrak{A}}$ of counting stars realized in $\mathfrak{A}$ is $\boldsymbol{\Upsilon}_{\mathrm{fin}} \,\dot\cup\, \boldsymbol{\Upsilon}_{\mathrm{inf}}$.

The kings. Let $\mathfrak{K} \subseteq \mathfrak{A}$, *the kings* of $\mathfrak{A}$, be the finite substructure with universe

$$K := \{a \in A \mid \mathrm{ctp}_{\mathfrak{A}}(a) \in \boldsymbol{\Upsilon}_{\mathrm{fin}}\}.$$

Formally we here admit that K can be empty and allow ourselves to talk about possibly empty structures. In any case, the kings are the elements of rare kinds. And as usual, they come with a court, which will consist of those elements that have special relationships with the kings.

The relationships with the kings are formalized by taking into account atomic 1-types with constant names for the kings. Recall that $\boldsymbol{\alpha}^K$ denotes the set of these types; $\boldsymbol{\alpha}^K$ is finite since K is finite.

Extended counting stars. The extended counting star is defined according to

$$\mathrm{ctp}^K_{\mathfrak{A}}(a) := \big(\mathrm{ctp}_{\mathfrak{A}}(a), \mathrm{atp}^K_{\mathfrak{A}}(a)\big).$$

Let $\boldsymbol{\Upsilon}^K$ stand for the finite set of all those elements of $\boldsymbol{\Upsilon} \times \boldsymbol{\alpha}^K$ that obey the following restriction:

$$(\gamma, \alpha) \in \boldsymbol{\Upsilon}^K \quad \text{if for all } \beta\text{: } \gamma(\beta) \geqslant \dot{\#}_{k \in K}\big(\alpha \models \beta[x,k]\big).$$

It is clear that any $\mathrm{ctp}^K_{\mathfrak{A}}(a) \in \boldsymbol{\Upsilon}^K$. Let Π be the natural projection

$$\begin{array}{rccc} \Pi\colon & \boldsymbol{\Upsilon}^K & \longrightarrow & \boldsymbol{\Upsilon} \\ & (\gamma,\alpha) & \longmapsto & \gamma. \end{array}$$

We also split $\boldsymbol{\Upsilon}^K$ according to the number of elements of $\mathfrak{A}$ that realize (γ,α) into

$$\boldsymbol{\Upsilon}^K = \boldsymbol{\Upsilon}^K_0 \,\dot\cup\, \boldsymbol{\Upsilon}^K_{\mathrm{fin}} \,\dot\cup\, \boldsymbol{\Upsilon}^K_{\mathrm{inf}}.$$

The court. Let the *court* be the substructure $\mathfrak{C} \subseteq \mathfrak{A}$ with universe

$$C := \{a \in A \mid \mathrm{ctp}^K_{\mathfrak{A}}(a) \in \boldsymbol{\Upsilon}^K_{\mathrm{fin}}\}.$$

It is clear that $\mathfrak{K} \subseteq \mathfrak{C} \subseteq \mathfrak{A}$, where $\mathfrak{K}$ and $\mathfrak{C}$ are finite and possibly empty. If we let $\boldsymbol{\Upsilon}^K_{\mathfrak{A}}$ be the set of $(\gamma,\alpha) \in \boldsymbol{\Upsilon}^K$ that are realized in $\mathfrak{A}$, then

$$\boldsymbol{\Upsilon}_{\mathfrak{A}} = \Pi\big(\boldsymbol{\Upsilon}^K_{\mathfrak{A}}\big) \qquad \text{and also} \qquad \boldsymbol{\Upsilon}_{\mathrm{inf}} = \Pi\big(\boldsymbol{\Upsilon}^K_{\mathrm{inf}}\big).$$

The latter assertion follows from the first one, if we notice that the fibres $\Pi^{-1}(\gamma)$ of the projection Π are finite.

The counting stars of kings only depend on their court and a very rough knowledge of the rest of the society.

Remark 3.1. *Given $\mathfrak{K}, \mathfrak{C}$ such that $\mathfrak{K} \subseteq \mathfrak{C}$ and $\boldsymbol{\Upsilon}^K_{\mathrm{inf}} \subseteq \boldsymbol{\Upsilon}^K$, it is possible to determine (recursively) $\mathrm{ctp}^K_{\mathfrak{A}}(k)$ for all $k \in K$.*

Proof. Let $\mathrm{ctp}^K_{\mathfrak{A}}(k) = (\gamma,\alpha)$. Then $\alpha = \mathrm{atp}^K_{\mathfrak{A}}(k) = \mathrm{atp}^K_{\mathfrak{K}}(k)$, and the interesting information to be reconstructed is $\gamma(\beta)$ for each $\beta \in \boldsymbol{\beta}$. We distinguish two cases: if there is some $(\gamma',\alpha') \in \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ such that $\alpha' \models \beta[k,x]$ then $\#_{a\in A}\big(\mathfrak{A} \models \beta[k,a]\big) = \omega$, so that $\gamma(\beta) = 2^+$. If there is no such $(\gamma',\alpha') \in \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$, then it must be the case that $\#_{a\in A}\big(\mathfrak{A} \models \beta[k,a]\big) = \#_{a \in C}\big(\mathfrak{A} \models \beta[k,a]\big)$, which may be determined in $\mathfrak{C}$. ⊣

Reduced stars. In some of the considerations to follow it will be important to classify elements of $A \setminus C$ according to how many β-edges they can have to elements *other than kings.* It is clear that this information can directly be extracted from $\mathrm{ctp}^K_{\mathfrak{A}}(a)$. For notational convenience let us introduce a function

$$\begin{array}{rrcl} \mathit{red}\colon & \boldsymbol{\Upsilon}^K & \longrightarrow & \boldsymbol{\Upsilon} \\ & (\gamma,\alpha) & \longmapsto & \gamma^- \qquad \text{where} \end{array}$$

$$\gamma^-(\beta) = \begin{cases} 2^+ & \text{if } \gamma(\beta) = 2^+ \\ \gamma(\beta) - \#_{k\in K}(\alpha \models \beta[x,k]) & \text{else.} \end{cases}$$

Observe that by the definition of $\boldsymbol{\Upsilon}^K$, $\mathit{red}(\gamma,\alpha)$ is a well-defined counting star — in particular no negative values can occur for $\gamma^-(\beta)$.

Characteristics. For infinite structures $\mathfrak{A}$ we abstract the following finite data as characteristic information, which we shall call the characteristic of $\mathfrak{A}$, or $\mathit{char}(\mathfrak{A})$ for short.

$$\boxed{\begin{array}{ll} \mathfrak{K} \subseteq \mathfrak{C} \quad \text{kings and court} & \\ F\colon C \;\to\; \boldsymbol{\Upsilon}^K & \\ \qquad c \;\mapsto\; \mathrm{ctp}^K_{\mathfrak{A}}(c) & \\ \boldsymbol{\Upsilon}^K_{\mathrm{inf}} \subseteq \boldsymbol{\Upsilon}^K & \end{array}} \qquad \mathit{char}(\mathfrak{A})$$

Let M be the recursive set of all tuples $(\mathfrak{K},\mathfrak{C},F,X)$, where $\mathfrak{K} \subseteq \mathfrak{C}$ are finite τ-structures (possibly empty), F is a mapping $F\colon C \to \boldsymbol{\Upsilon}^K$ and $\emptyset \neq X \subseteq \boldsymbol{\Upsilon}^K$.

Consider the situation in which rather than $\mathfrak{A}$ itself, only $\mathit{char}(\mathfrak{A})$ is presented and we want to know whether $\mathfrak{A} \models \varphi$ for φ in normal form. By Lemma 2.4 it suffices to determine $\boldsymbol{\Upsilon}_{\mathfrak{A}} = \boldsymbol{\Upsilon}_{\mathrm{fin}} \cup \boldsymbol{\Upsilon}_{\mathrm{inf}}$. But $\boldsymbol{\Upsilon}_{\mathrm{fin}} = \{\mathrm{ctp}_{\mathfrak{A}}(k) \mid k \in K\}$ may, by Remark 3.1, effectively be determined from the knowledge of $\mathfrak{K}$, $\mathfrak{C}$, and $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$. $\boldsymbol{\Upsilon}_{\mathrm{inf}}$ is just the projection $\Pi(\boldsymbol{\Upsilon}^K_{\mathrm{inf}})$. This yields the following.

Remark 3.2. *Given $\mathit{char}(\mathfrak{A})$ and φ in normal form, it can be decided whether $\mathfrak{A} \models \varphi$.*

It follows that the set of C^2-sentences in normal form that are in *inf-sat*(C^2) is r.e., provided that we can show that $\{\mathit{char}(\mathfrak{A}) \mid \mathfrak{A} \text{ an infinite } \tau\text{-structure}\}$ is a recursive subset of M. This is shown in the following section.

4. Decidability of the characteristics

Theorem 4.1. *Given $(\mathfrak{K},\mathfrak{C},F,X) \in M$, it is decidable whether there is an infinite $\mathfrak{A}$ such that $(\mathfrak{K},\mathfrak{C},F,X) = \mathit{char}(\mathfrak{A})$.*

The proof is separated into two parts: in the first step we isolate three necessary conditions; these are shown to be sufficient in the second step.

4.1. Three necessary conditions

Let $(\mathfrak{K},\mathfrak{C},F,X) = \mathit{char}(\mathfrak{A})$ for some infinite $\mathfrak{A}$. We may write $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ and $\boldsymbol{\Upsilon}_{\mathrm{inf}}$ for X and its projection $\Pi(X)$. Then the following conditions C1–C3 are satisfied.

Condition C1: Compatibility of F with $\mathfrak{K}$ and $\mathfrak{C}$. C1 actually is a group of rather simple conditions. They assure for vertices in $\mathfrak{C}$ that F specifies atomic types (with parameters in K) in accordance with the actual atomic types in $\mathfrak{C}$; and that the counting stars specified by F are such that (1) no vertex already has more outgoing β-edges within $\mathfrak{C}$ than are allowed by its counting star, and (2) if a vertex has fewer outgoing β-edges within $\mathfrak{C}$ than required by its counting star, then there are extended counting stars in $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ which accept incoming β-edges.

C1(a): for all $k \in \mathfrak{K}$: if $F(k) = (\gamma_0,\alpha_0)$, then

(i) $\gamma_0 \notin \boldsymbol{\Upsilon}_{\mathrm{inf}}$.

(ii) $\alpha_0 = \mathrm{atp}^K_{\mathfrak{K}}(k)$.

(iii) for all $\beta \in \boldsymbol{\beta}$:

- if $\gamma_0(\beta) \in \{0,1\}$, then $\gamma_0(\beta) = \#_{c\in C}\big(\mathrm{atp}_{\mathfrak{C}}(k,c) = \beta\big)$ and for all $(\gamma,\alpha) \in \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$: $\alpha \models \neg\overline{\beta}[x,k]$.
- if $\gamma_0(\beta) = 2^+$ and $\#_{c\in C}\big(\mathrm{atp}_{\mathfrak{C}}(k,c) = \beta\big) < 2$, then there is some $(\gamma,\alpha) \in \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ such that $\alpha \models \overline{\beta}[x,k]$.

C1(b): for all $c \in \mathfrak{C} \setminus \mathfrak{K}$: if $F(c) = (\gamma_0,\alpha_0)$, then

(i) $(\gamma_0,\alpha_0) \notin \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$, but $\gamma_0 \in \boldsymbol{\Upsilon}_{\mathrm{inf}}$.

(ii) $\alpha_0 = \mathrm{atp}^K_{\mathfrak{C}}(c)$.

(iii) for all $\beta \in \boldsymbol{\beta}$:

- $\gamma_0(\beta) \geqslant \#_{c'\in C}\big(\mathrm{atp}_{\mathfrak{C}}(c,c') = \beta\big)$.
- if $\gamma_0(\beta) > \#_{c'\in C}\big(\mathrm{atp}_{\mathfrak{C}}(c,c') = \beta\big)$, then there is some $(\gamma,\alpha) \in \boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ whose reduced counting star $\gamma^- := \mathit{red}(\gamma,\alpha)$ satisfies $\gamma^-(\overline{\beta}) > 0$.

Proof of necessity. In both cases (i) and (ii) are obvious. For (iii) in (a) observe that $\#_{a\in A}\big(\mathrm{atp}_{\mathfrak{A}}(k,a)=\beta\big)$ is infinite iff there is a $(\gamma,\alpha)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}\,:\,\alpha\models\overline{\beta}[x,k]$. The implication from left to right uses finiteness of $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ and the fact that any a with $\mathrm{ctp}^K_{\mathfrak{A}}(a)\notin\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ belongs to the finite $\mathfrak{C}$.

If $\#_{a\in A}\big(\mathrm{atp}_{\mathfrak{A}}(k,a)=\beta\big)$ is finite (which in particular must be the case if $\gamma_0(\beta)=0,1$), then $\#_{a\in A}\big(\mathrm{atp}_{\mathfrak{A}}(k,a)=\beta\big)=\#_{a\in C}\big(\mathrm{atp}_{\mathfrak{C}}(k,a)=\beta\big)$.

For necessity of (b)(iii), observe that, if $\gamma_0(\beta)>\#_{c'\in C}\big(\mathrm{atp}_{\mathfrak{C}}(c,c')=\beta\big)$, then there must be an $a\in A\setminus C$ such that $\mathrm{atp}_{\mathfrak{A}}(c,a)=\beta$. But $(\gamma,\alpha):=\mathrm{ctp}^K_{\mathfrak{A}}(a)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ and, as $c\notin K$, the reduced counting star $\gamma^-=\mathit{red}(\gamma,\alpha)$ admits the incoming β-edge (c,a), whence $\gamma^-(\overline{\beta})>0$. ⊣

Condition C2: A closure property of $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$. This condition assures that $\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ is *closed* in the sense that for any extended star type in the infinite part that requires outgoing β-edges to elements other than kings, there is another extended star type in the infinite part that can receive incoming β-edges from elements other than kings.

C2: For all $(\gamma_0,\alpha_0)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ and $\beta\in\boldsymbol{\beta}$: if $\gamma_0(\beta)>\#_{k\in K}\big(\alpha_0\models\beta[x,k]\big)$, then there is some $(\gamma,\alpha)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ whose reduced counting star $\gamma^-:=\mathit{red}(\gamma,\alpha)$ satisfies $\gamma^-(\overline{\beta})>0$.

Proof of necessity. Consider all elements $b\in A\setminus C$ with $\mathrm{ctp}^K_{\mathfrak{A}}(b)=(\gamma_0,\alpha_0)$ — there are infinitely many of them. $\gamma_0(\beta)>\#_{k\in K}\big(\alpha_0\models\beta[x,k]\big)$ implies that for each such b there is some $a\in A\setminus K$ for which $\mathrm{atp}_{\mathfrak{A}}(b,a)=\beta$. If for some b there even is such an $a\in A\setminus C$, then $(\gamma,\alpha):=\mathrm{ctp}^K_{\mathfrak{A}}(a)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$ is as desired.

If on the other hand for all b there are only $a\in C\setminus K$ with $\mathrm{atp}_{\mathfrak{A}}(b,a)=\beta$, then for some of these finitely many a it must be the case that $\gamma:=\mathrm{ctp}_{\mathfrak{A}}(a)\in\boldsymbol{\Upsilon}_{\mathrm{inf}}$ has $\gamma(\overline{\beta})=2^+$. Let α be such that $(\gamma,\alpha)\in\boldsymbol{\Upsilon}^K_{\mathrm{inf}}$. As $\gamma(\overline{\beta})=2^+$, also the reduced $\gamma^-=\mathit{red}(\gamma,\alpha)$ has $\gamma^-(\overline{\beta})=2^+$, and (γ,α) is as desired. ⊣

Condition C3: A homogeneity property of $\boldsymbol{\Upsilon}_{\mathrm{inf}}$. This last condition asserts that for all pairs of distinct vertices in the infinite part there is at least one β to connect them which is not limited (i.e., has value 2^+) according to the specified star types at either end.

C3: For all $\gamma,\gamma'\in\boldsymbol{\Upsilon}_{\mathrm{inf}}$ there is some $\beta\in\boldsymbol{\beta}$ such that $\gamma(\beta)=\gamma'(\overline{\beta})=2^+$.

This uses a result from Ramsey theory, see Theorem 1 in Chapter 5 of [6].

Theorem 4.2. *If the edges of the complete bipartite graph $K_{\omega,\omega}$ are coloured with finitely many colours, then it contains monochromatically coloured copies of the complete bipartite graph $K_{n,n}$, for all n.*

Proof of necessity of C3. Note that C3 applies even to the case $\gamma=\gamma'$. Embed $K_{\omega,\omega}$ injectively into $A\setminus K$ in such a way that the parts (of the bipartition) are mapped into $\{a\in A\mid \mathrm{ctp}_{\mathfrak{A}}(a)=\gamma\}$ and $\{a\in A\mid \mathrm{ctp}_{\mathfrak{A}}(a)=\gamma'\}$, respectively. The $\beta\in\boldsymbol{\beta}$ induce a finite colouring of the edges, and any β that admits a β-coloured copy of $K_{2,2}$ is as desired. ⊣

Remark 4.3. *Conditions C1–C3 are recursive for given $(\mathfrak{K},\mathfrak{C},F,X)\in M$.*

4.2. Sufficiency of C1–C3

It remains to prove the following.

Proposition 4.4. *For any $(\mathfrak{K},\mathfrak{C},F,X)\in M$ that satisfies C1–C3, there is an infinite $\mathfrak{A}$ such that $(\mathfrak{K},\mathfrak{C},F,X)=\mathit{char}(\mathfrak{A})$.*

Proof. Let $(\mathfrak{K},\mathfrak{C},F,X)$ satisfying C1–C3 be given. Beyond the elements of $\mathfrak{C}$, the desired structure $\mathfrak{A}$ has to have, for each of the finitely many $(\gamma,\alpha)\in X$, an infinite sequence of vertices that realize exactly that (γ,α). We therefore put $A:=C\,\dot\cup\,(\omega\times X)$ for the universe of $\mathfrak{A}$. Think of the new vertices in $V:=\omega\times X$ as divided into finitely many infinite boxes, according to their second components, which specify the extended star types these vertices shall eventually realize.

The task now is to declare atomic types for all pairs of vertices of A in such a way that

(a) a consistent interpretation of a τ-structure $\mathfrak{A}$ over A with $\mathfrak{K}\subseteq\mathfrak{C}\subseteq\mathfrak{A}$ is obtained,

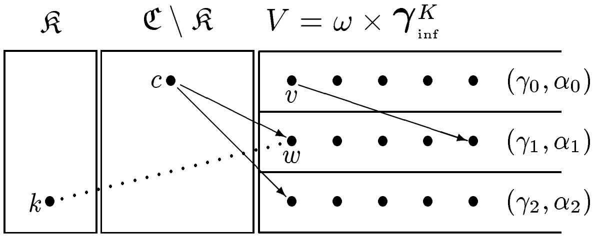
(b) all elements $c\in C$ satisfy the counting star and atomic τ_K-type prescribed by F: $\mathrm{ctp}^K_{\mathfrak{A}}(c)=F(c)$,

(c) all elements of V satisfy the counting star and atomic τ_K-type suggested by their X-component: $\mathrm{ctp}^K_{\mathfrak{A}}(v)=(\gamma,\alpha)$ for $v=\big(m,(\gamma,\alpha)\big)$, $m\in\omega$.

With respect to (c), let $\widehat{F}$ be the extension of the given F to all of A according to $\widehat{F}\big(m,(\gamma,\alpha)\big):=(\gamma,\alpha)$. Then (b) and (c) require that $\mathrm{ctp}^K_{\mathfrak{A}}(a)=\widehat{F}(a)$ for all $a\in A$. Indeed $\mathit{char}(\mathfrak{A})=\big(\mathfrak{K},\mathfrak{C},F,X\big)$ then follows. In particular it should be noted that (b) and (c) imply that $\mathfrak{K}$ and $\mathfrak{C}$ really become the kings and court of $\mathfrak{A}$.

Giving a full interpretation as a τ-structure to $\mathfrak{A}$ should be thought of as allocating atomic 2-types $\beta\in\boldsymbol{\beta}$ to any pair of distinct vertices in A. Thus one may think of successively *putting β-edges for suitable β* between any two distinct vertices. Note that atomic 1-types get settled automatically through the allocation of atomic 2-types.

In the following we first give a rough and intuitive sketch of how (a)–(c) can be achieved, and afterwards a more detailed and definite description. It may be useful to consult the following diagram, which indicates the scenario we start with and also highlights some aspects of the tasks discussed below.



Edges involving two vertices from C are completely specified in $\mathfrak{C}$, of course. Observe also that all edges between K and V are forced by requirement (c): $\text{atp}^K_{\mathfrak{A}}(w) \models \beta[k, x]$ for exactly one $\beta \in \boldsymbol{\beta}$, and $\text{atp}^K_{\mathfrak{A}}(w)$ is specified by the box that w belongs to.

In fact it suffices to satisfy (b) for the elements of $C \setminus K$, because (c) then implies that (b) is also fulfilled for all $k \in K$. This follows by Remark 3.1 together with condition C1(a): if F specifies the counting star of $k \in K$ to be γ and if $\gamma(\beta) = 0$ or 1, then C1(a) guarantees that (1) the corresponding number of β-edges is already put right in $\mathfrak{C}$ and that (2) no more β-edges from k to elements of V can be introduced if the $\text{atp}^K_{\mathfrak{A}}(v)$ are put right according to specification. If on the other hand $\gamma(\beta) = 2^+$, then either there are already at least two β-edges from k to elements of $\mathfrak{C}$, or condition C1(a) (iii) implies that infinitely many more will be introduced if the $\text{atp}^K_{\mathfrak{A}}(v)$ for $v \in V$ are settled according to specification.

It is important to observe that as yet, i.e., with β-edges attributed to pairs over C according to the given $\mathfrak{C}$, no vertex has more outgoing β-edges than it is meant to have according to $\widehat{F}$. This follows from condition C1.

> *In allocating new edges, this property must be preserved at both ends of any new edge!*

With this provision in mind, the remaining tasks will be settled in the following order:

T1: introduction of sufficiently many outgoing β-edges of respective kinds at each $c \in C \setminus K$ to get (b) right.

T2: introduction of sufficiently many outgoing β-edges of respective kinds at each $v \in V$ to get (c) right.

T3: declaration of all remaining edges between pairs of distinct vertices (without affecting their prospective counting stars any more).

A rough sketch. Compare the diagram above for the argument. The outgoing edges for T1 and T2 (we only talk of those that do not go to kings by now) can all be chosen to go to vertices in V (rather than possibly to $C \setminus K$). Consider T1 for $c \in C \setminus K$. If this c requires β-edges to vertices outside C, then condition C1(b) (iii) says that there is some box in the infinite part, whose vertices w can accept an incoming β-edge; this precisely means that if (γ_1, α_1) is the specification for w, then the reduced star type γ^- of w has $\gamma^-(\beta) > 0$ — note that γ^- rather than γ has to be considered as α_1 may already specify incoming β-edges from kings (the dotted line in the diagram could be a β-edge). Similarly for some $v \in V$ that requires β-edges to vertices other than kings: condition C2 now guarantees that such edges can be directed to vertices in V, whose reduced star type lets them accept incoming β-edges.

All edges introduced in phases T1 and T2 can be chosen independently in the sense that no two such edges ever go to the same vertex in V — thereby preventing the danger that any individual $v \in V$ gets overloaded. (The right order for going through the $v \in V$ for T2 is to treat them according to increasing first component, or column-wise in the diagram.)

At the end of this phase, all vertices have a correct number of outgoing β-edges for all β. Their counting stars would be all right, only there remain edges connecting elements in V and between elements in $C \setminus K$ and V to be declared in phase T3. This is where condition C3 becomes essential, as it guarantees edges β that may be used between any pair of distinct vertices from $A \setminus K$ without affecting the count that is taken at either end: simply because multiplicity 2^+ cannot be spoilt by the introduction of extra edges of that kind.

The explicit construction. A detailed strategy to achieve T1–T3 can actually be given by means of choice functions telling into which one of the infinite compartments of V, $\omega \times \{(\gamma, \alpha)\}$, edges are to be directed during stages T1 and T2, and which β-edges to choose in T3. Such choice functions, f, g and h may be fixed as follows. We already write $\boldsymbol{\gamma}^K_{\text{inf}}$ and $\boldsymbol{\gamma}_{\text{inf}}$ for the given $X \subseteq \boldsymbol{\gamma}^K$ and its projection $\Pi(X)$, since this is what we want these sets to become.

- $$f : (C \setminus K) \times \boldsymbol{\beta} \longrightarrow \boldsymbol{\gamma}^K_{\text{inf}},$$

such that if $F(c) = (\gamma_0, \alpha_0)$ and $\gamma_0(\beta) > \#_{c' \in C}\big(\text{atp}_{\mathfrak{C}}(c, c') \models \beta\big)$, then $(\gamma, \alpha) := f(c, \beta)$ is as guaranteed in C1(b), i.e., for the reduced counting star $\gamma^- := red(\gamma, \alpha)$: $\gamma^-(\overline{\beta}) > 0$.

- $$g : \boldsymbol{\Upsilon}^K_{\text{inf}} \times \boldsymbol{\beta} \longrightarrow \boldsymbol{\Upsilon}^K_{\text{inf}},$$

such that if $\gamma_0(\beta) > \#_{k \in K}\big(\alpha_0 \models \beta[x,k]\big)$, then $(\gamma, \alpha) := g\big((\gamma_0, \alpha_0), \beta\big)$ is as guaranteed in C2, i.e., for the reduced counting star $\gamma^- := red(\gamma, \alpha)$: $\gamma^-(\overline{\beta}) > 0$.

- $$h : \boldsymbol{\Upsilon}_{\text{inf}} \times \boldsymbol{\Upsilon}_{\text{inf}} \longrightarrow \boldsymbol{\beta},$$

such that $\beta := h\big(\gamma, \gamma'\big)$ is as guaranteed by C3, i.e., $\gamma(\beta) = \gamma'(\overline{\beta}) = 2^+$.

With these choice functions, we can give a definite description of the desired $\mathfrak{A}$, or of the allocation of edges β to all remaining pairs of distinct elements, according to T1–T3 above. To be quite definite about the sequence in which tasks are settled, let $C \setminus K = \{c_1, \ldots, c_r\}$, $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_s\}$, and $\boldsymbol{\Upsilon}^K_{\text{inf}} = \{(\gamma_1, \alpha_1), \ldots, (\gamma_t, \alpha_t)\}$ be enumerations of the respective sets without repetitions.

T1: For $i = 1, \ldots, r$ / for $j = 1, \ldots, s$:
if $\Pi\big(F(c_i)\big) = \gamma$, and if $\gamma(\beta_j)$ exceeds the number of outgoing β_j-edges that c_i already has, then choose m minimal in ω such that $v = (m, f(c_i, \beta_j)) \in V$ does not yet have any incoming edges from vertices in $C \setminus K$, and put a β_j-edge, i.e., make $\text{atp}_{\mathfrak{A}}(c_i, v) = \beta_j$ for that $v \in V$. This procedure is carried out one or two times, depending on whether $\gamma(\beta_j)$ exceeds the number of outgoing β_j-edges that c originally has within $\mathfrak{C}$ by 1 or 2.

This is compatible with the requirements on $\text{ctp}^K(v)$, since f is such that $\text{ctp}^K(v)$ admits at least one $\overline{\beta}_j$-edge to vertices outside K.

T2: For $n = 0, 1, 2, \ldots$ / for $i = 1, \ldots, t$ / for $j = 1, \ldots, s$:
if $\gamma_i(\beta_j)$ exceeds the number of outgoing β_j-edges that $v = \big(n, (\gamma_i, \alpha_i)\big)$ already has, choose m minimal in ω such that $w = \big(m, g\big((\gamma_i, \alpha_i), \beta_j\big)\big) \in V$ does not yet have any incoming edges from outside K, and put a β_j-edge. That is, put $\text{atp}_{\mathfrak{A}}(v, w) = \beta_j$ for that $w \in V$. Again, this procedure may have to be applied once or twice to the same v, depending on the number of β_j-edges required.

Compatibility with the requirements on $\text{ctp}^K(w)$ is guaranteed by the choice of g.

T3: It remains to settle all remaining atomic 2-types, namely those between two distinct vertices v and v' in V, or between $v \in V$ and $c \in C \setminus K$, that have not yet been connected by any β-edge.
For definiteness let $v = \big(m, (\gamma_i, \alpha_i)\big)$ and $v' = \big(m', (\gamma_{i'}, \alpha_{i'})\big)$ with $i < i'$ or $m < m'$ in the first case, and $v = \big(m, (\gamma_i, \alpha_i)\big)$ and $\gamma_{i'} := \Pi\big(F(c)\big)$ in the second case. We then put $\text{atp}_{\mathfrak{A}}(v, v') = \beta$ respectively $\text{atp}_{\mathfrak{A}}(v, c) = \beta$, for $\beta := h(\gamma_i, \gamma_{i'})$.

Compatibility with $\text{ctp}(v) := \gamma$ and $\text{ctp}(v') := \gamma'$ or $\text{ctp}(c) := \gamma'$ is clear, since the β selected by h is such that $\gamma(\beta) = \gamma'(\overline{\beta}) = 2^+$.

This finishes the proof of sufficiency: $\mathfrak{A}$ as constructed has $char(\mathfrak{A}) = \big(\mathfrak{K}, \mathfrak{C}, F, X\big)$. $\dashv$

This also completes the proof of Theorem 1.2.

5. Finite satisfiability

We now sketch the proof that $\mathit{fin\text{-}sat}(C^2)$ is decidable. The structure of the proof is quite similar to the general case, although the combinatorics are more complex. Most of the same terminology will be used, but in some cases the definition of a term is changed. For example, kings were defined above as those elements that realize a counting star that is realized only finitely many times. In the finite case, kings will be those elements that realize a counting star that is realized only a 'small' number of times.

The major difference between the general case and the finite case arises because we have to analyze carefully the precise number of times any 'large' counting star is realized. In the general case, all the large types were realized ω many times. In the finite case, the numbers of elements that realize each of the large counting stars have to satisfy a Boolean combination of linear equations and inequalities. The most important of these equations say something, for some atomic 2-type β, about the relationship between the number of $a \in A$ such that $\exists x \beta[a, x]$ and the number of $a \in A$ such that $\exists x \beta[x, a]$, e.g., that these numbers are equal.

The main new idea involves giving a rather different definition of the analog of the court. We replace the court by a set of *nobles*, which are those elements that are either kings or (roughly) are incident to a β-edge that does not get realized many times in the non-regal part of the structure.

6. Analysis of finite structures

Before defining the set of kings, we need the finite version of the bipartite Ramsey theorem. Again, see Theorem 1 in Chapter 5 of [6].

Theorem 6.1. *For all natural numbers p and t, there is an m such that if the edges of $K_{m,m}$ are t-coloured, then there exists a monochromatic $K_{p,p}$.*

Let $R(p,t)$ denote the least m with this property.

Let $\mathfrak{A}$ be a fixed finite τ-structure. We define $r = 2 \cdot R(2, |\boldsymbol{\beta}|) + 16|\boldsymbol{\beta}|^2$. (The exact value of r does not matter; it just has to be large enough for a number of different purposes.) Split $\boldsymbol{\gamma}$ into three disjoint subsets

$$\boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \,\dot{\cup}\, \boldsymbol{\gamma}_{\mathrm{sm}} \,\dot{\cup}\, \boldsymbol{\gamma}_{\mathrm{lg}}$$

according to whether the number of times that γ is realized in $\mathfrak{A}$ is 0, is $\geqslant 1$ and $< r$, or is $\geqslant r$.

The kings and the nobles. Let $\mathfrak{K} \subseteq \mathfrak{A}$, the *kings* of $\mathfrak{A}$, be the finite substructure with universe

$$K := \{a \in A \mid \mathrm{ctp}_{\mathfrak{A}}(a) \in \boldsymbol{\gamma}_{\mathrm{sm}}\}.$$

The role of the kings in the finite case is exactly like their role in the general case. Observe that for all $a_0, a_1 \in A \setminus K, \gamma_i = \mathrm{ctp}_{\mathfrak{A}}(a_i)$, there is a β such that $\gamma_0(\beta) = \gamma_1(\overline{\beta}) = 2^+$.

Before introducing the nobles, we need the following definitions. For a fixed structure $\mathfrak{A}$, let $\boldsymbol{\gamma}^K_{\mathrm{lg}}$ be the set of extended counting stars that are realized in $\mathfrak{A}$ by a non-king. Below, when $\gamma^K = (\gamma_0, \alpha_0)$ is an extended counting star, we shall occasionally write $\gamma^K(\beta)$ for $\gamma_0(\beta)$. Also, when 2^+ occurs in an equation, it is to be evaluated as 2. For example $2^+ > 2$ is false, $2^+ - 1 = 1$, and $2^+ - 0 = 2^+$. For $\mathfrak{A}$ and $\beta \in \boldsymbol{\beta}$, let $\Lambda(\beta)$ denote the set

$$\{\gamma^K := (\gamma_0, \alpha_0) \in \boldsymbol{\gamma}^K_{\mathrm{lg}} \mid \gamma_0(\beta) > \#_{k \in K}(\alpha_0 \models \beta[x,k])\}.$$

That is, $\gamma^K \in \Lambda(\beta)$ if for every $a \in A \setminus K$ that realizes γ^K, there 'must' be an $a' \in A \setminus K$ such that $\beta(a, a')$. We define $\boldsymbol{\beta}_{\mathrm{lg}} = \{\beta \mid \Lambda(\beta) \neq \emptyset\}$.

The *noble elements* of $\mathfrak{A}$ are specified in terms of *noble edges*, which we now describe. Intuitively, an edge is noble if it is an atomic 2-type that does not get realized too many times by pairs of non-kings, but the exact definition is more subtle and requires considering a number of different cases. Here we only discuss one. Define an edge $\beta \in \boldsymbol{\beta}_{\mathrm{lg}}$ to be *separate* (in a structure $\mathfrak{A}$) iff $\Lambda(\beta) \cap \Lambda(\overline{\beta}) = \emptyset$. If $\beta \in \boldsymbol{\beta}_{\mathrm{lg}}$ is separate, and for all $\gamma^K \in \boldsymbol{\gamma}^K_{\mathrm{lg}} : \gamma^K(\beta), \gamma^K(\overline{\beta}) \neq 2^+$, then it is *noble*, $\beta \in \boldsymbol{\beta}_N$, iff $|\{a \in A \setminus K \mid \mathrm{ctp}^K_{\mathfrak{A}}(a) \in \Lambda(\beta)\}| < 4|\boldsymbol{\beta}|$.

Definition 6.2. *For all $a \in \mathfrak{A}$, a is a* noble *iff $a \in K$ or there is a $\beta \in \boldsymbol{\beta}_N$ such that $\mathrm{ctp}^K_{\mathfrak{A}}(a) \in \Lambda(\beta)$.*

Villagers are those elements that are not nobles. We split $\boldsymbol{\gamma}^K$ into

$$\boldsymbol{\gamma}^K = \boldsymbol{\gamma}^K_0 \,\dot{\cup}\, \boldsymbol{\gamma}^K_N \,\dot{\cup}\, \boldsymbol{\gamma}^K_V$$

according to whether γ^K is not realized, is realized by nobles, or by villagers. The set of *villager edges*, $\boldsymbol{\beta}_V$, contains exactly those $\beta \in \boldsymbol{\beta}_{\mathrm{lg}} - \boldsymbol{\beta}_N$.

Finite characteristics. Of course, unlike infinite structures, every finite structure can be described in a finite way. But to show that *fin-sat*(C^2) is decidable, we want there to be a finite number of finite characteristics $\mathcal{H}$ such that it is decidable whether there is a finite structure $\mathfrak{A}$ such that *fchar*$(\mathfrak{A}) = \mathcal{H}$. Otherwise the basic idea of *fchar*$(\mathfrak{A})$ is quite similar to that of *char*$(\mathfrak{A})$. One difference is that instead of a substructure $\mathfrak{C}$, there will be a weak substructure $\mathfrak{N}'$, explained below. The finite characteristic of a finite structure consists of the following information.

$$\boxed{\begin{array}{lll} \mathfrak{K} \subseteq \mathfrak{N}' & \text{kings and nobles}' & \\ F : N' & \to & \boldsymbol{\gamma}^K \\ \quad a & \mapsto & \mathrm{ctp}^K_{\mathfrak{A}}(a) \\ \boldsymbol{\gamma}^K_N \text{ and } \boldsymbol{\gamma}^K_V & & \end{array}} \qquad \textit{fchar}(\mathfrak{A})$$

We now define precisely the structure $\mathfrak{N}'$. The idea is that $\mathfrak{N}'$ is an incomplete version of $\mathfrak{N}$, obtained by 'removing' all edges connecting non-regal nobles. (This will provide extra freedom when we construct structures from *fchar*'s.) The universe of $\mathfrak{N}'$ is identical to that of $\mathfrak{N}$, and for all $a \in N', \mathrm{atp}^K_{\mathfrak{N}'}(a) = \mathrm{atp}^K_{\mathfrak{N}}(a)$. Also, for all pairs $a_0, a_1 \in N'$ such that at least one of the elements is in K, $\mathrm{atp}_{\mathfrak{N}'}(a_0, a_1) = \mathrm{atp}_{\mathfrak{N}}(a_0, a_1)$. But if both a_0 and a_1 are in $N' \setminus K$ (and distinct), then for all binary $Rxy \in \tau$, $\mathfrak{N}' \models \neg R a_0 a_1$.

The following analog of Remark 3.2 is proved in the same way.

Remark 6.3. *Given* fchar$(\mathfrak{A})$ *and φ in normal form, it can be decided whether $\mathfrak{A} \models \varphi$.*

Observe that there are only finitely many *fchar*'s, since the size of $\mathfrak{N}'$ is bounded by some n, a function of $|\boldsymbol{\alpha}|$ and $|\boldsymbol{\beta}|$. Let M be the *(finite!)* set of all tuples $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ where $\mathfrak{K}$ and $\mathfrak{N}'$ are finite τ-structures, $|N'| \leqslant n$, F is a mapping $F : N' \to \boldsymbol{\gamma}^K$ and $X_N, X_V \subseteq \boldsymbol{\gamma}^K$ are disjoint subsets.

It only remains to show that there are decidable necessary and sufficient conditions for a tuple $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ to be the *fchar* of a finite structure. The first three conditions will essentially be conditions C1–C3 from above. The new conditions, taken together, assert that there is a solution to a set of linear equations and inequalities, where each variable represents the number of times that an extended counting star is realized. The conditions are decidable, and it will be easy to see that they are also necessary.

Some simple combinatorial lemmas are used in the proof of sufficiency. The following one is representative.

Lemma 6.4. *(i) For all $d, m \in \omega$, if $m \geqslant 2d$, $G = (V_1, V_2, E)$ is a bipartite graph of degree $\leqslant d$, and $|V_1| = |V_2| = m$, then there exists a bijection $f : V_1 \longrightarrow V_2$ such that for all $v \in V_1, (v, f(v)) \notin E$.*

(ii) Furthermore, under the above hypothesis, if $c \leqslant m - 2d$, and g is an injective partial function from V_1 to V_2, such that for all $v \in dom(g), (v, g(v)) \notin E$ and $|dom(g)| \leqslant c$, then there is a function f, as required above, that extends g.

7. Decidability of the characteristics

Theorem 7.1. *Given $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V) \in M$, it is decidable whether there is a finite $\mathfrak{A}$ such that $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V) = fchar(\mathfrak{A})$.*

Again, we first isolate some necessary conditions, and then show that they are also sufficient.

7.1. The necessary conditions

Let $(\mathfrak{K}, \mathfrak{C}, F, X_N, X_V) = fchar(\mathfrak{A})$ for some finite $\mathfrak{A}$. We may write $\boldsymbol{\Upsilon}_N^K$ and $\boldsymbol{\Upsilon}_V^K$ for X_N and X_V. Likewise, we may write $\boldsymbol{\Upsilon}_{\mathrm{lg}}^K, \Lambda_{\boldsymbol{\beta}}, \boldsymbol{\beta}_N$, and $\boldsymbol{\beta}_V$, since it is easy to see that these are determined by the above information. Then the following conditions are satisfied. Conditions C1–C3 are close to the corresponding conditions in the proof of the general result. In this version, we only state the most interesting new conditions, C4 and C5.

Condition C4: Realizability of nobles. This condition says, roughly, that $(\mathfrak{K}, \mathfrak{N}', F)$ is compatible with the nobles of a structure with $fchar = (\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$.

C4: There is a structure $\mathfrak{N}, N = N'$, that satisfies:

(i) For all $a \in N, \mathrm{atp}_{\mathfrak{N}'}(a) = \mathrm{atp}_{\mathfrak{N}}(a)$.

(ii) For all $a_0 \in K$, and all $a_1 \in N, \mathrm{atp}_{\mathfrak{N}'}(a_0, a_1) = \mathrm{atp}_{\mathfrak{N}}(a_0, a_1)$.

(iii) For all $a \in N, F(a) = (\gamma, \alpha)$, and all $\beta \in \boldsymbol{\beta}, \#_{a' \in N}(\mathfrak{N}' \models \beta(a, a')) \leqslant \gamma(\beta)$. Furthermore, if $\beta \in \boldsymbol{\beta}_N$, then $\#_{a' \in N}(\mathfrak{N}' \models \beta[a, a']) = \gamma(\beta)$.

Condition C5: Correlating $\Lambda(\beta)$ and $\Lambda(\overline{\beta})$. This condition consists of a set of linear constraints, indexed by $\beta \in \boldsymbol{\beta}_V$, that relate the number of elements that realize a type in $\Lambda(\beta)$ with the number of elements that realize a type in $\Lambda(\overline{\beta})$. For each $\gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K$, let $n(\gamma^K) = |\{a \in A \mid \mathrm{ctp}_{\mathfrak{A}}^K(a) = \gamma^K\}|$. And for $\beta \in \boldsymbol{\beta}_V$, let $m(\beta) = \sum_{\gamma^K \in \boldsymbol{\Lambda}(\beta)} n(\gamma^K)$.

For each $\beta \in \boldsymbol{\beta}_V$, let $H_\beta(x) : \boldsymbol{\Upsilon}_{\mathrm{lg}}^K \longrightarrow \{0, 1, 2^+\}$ be the function $H_\beta((\gamma, \alpha)) = \gamma(\beta) \dot{-} \#_{k \in K}(\alpha \models \beta[x, k])$. That is, $H_\beta((\gamma, \alpha))$ is the minimum number of outgoing β-edges *in* $A \setminus K$ that an element that realizes (γ, α) must have. For each $\beta \in \boldsymbol{\beta}_V$, we add the condition

C5(β)**:**

(i) if for all $\gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K, \gamma^K(\beta), \gamma^K(\overline{\beta}) \neq 2^+$ then

$$\sum_{\gamma^K \in \boldsymbol{\Lambda}(\beta)} n(\gamma^K) = \sum_{\gamma^K \in \boldsymbol{\Lambda}(\overline{\beta})} n(\gamma^K).$$

(ii) if β is symmetric and for all $\gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K$, $\gamma^K(\beta) \neq 2^+$, we add the requirement that $m(\beta)$ is even.

(iii) if there is a $\gamma_0^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K$ such that $\gamma_0^K(\beta) = 2^+$ and for all $\gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K, \gamma^K(\overline{\beta}) \neq 2^+$, then

$$\sum_{\gamma^K \in \boldsymbol{\Lambda}(\beta)} H_\beta(\gamma^K) \cdot n(\gamma^K) \leqslant \sum_{\gamma^K \in \boldsymbol{\Lambda}(\overline{\beta})} n(\gamma^K).$$

The sum on the left is the minimum number of β-edges (connecting non-kings) that a structure must have in order to realize each $\gamma^K \in \Lambda_\beta, n(\gamma^K)$ many times.

Remark 7.2. *Given $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$, it is decidable whether there are numbers $\langle n(\gamma^K) \mid \gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K\rangle$ such that $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ and $\langle n(\gamma^K) \mid \gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K\rangle$ satisfy the conditions.*

C1–C3 are simply as in the general case. The decidability of C4 is easy. Together, the other conditions assert that there is a solution in the natural numbers to a (positive) Boolean combination of linear equations and inequalities. This can be expressed in a first-order way, using only the function $x + y$. Therefore, by the decidability of the theory of Presburger arithmetic, it is decidable whether there is a simultaneous solution to these equations.

7.2. Sufficiency of the conditions

It remains to prove the following.

Proposition 7.3. *For any $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ and $\langle n(\gamma^K) \mid \gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K\rangle$ that satisfy the above conditions, there is a finite structure $\mathfrak{A}$ such that $fchar(\mathfrak{A}) = (\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ and for every $\gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K$, $n(\gamma^K) = |\{a \in A \mid \mathrm{ctp}_{\mathfrak{A}}^K(a) = \gamma^K\}|$.*

Sketch of proof. Let $(\mathfrak{K}, \mathfrak{N}', F, X_N, X_V)$ and $\langle n(\gamma^K) \mid \gamma^K \in \boldsymbol{\Upsilon}_{\mathrm{lg}}^K\rangle$ be given. Beyond the elements of $\mathfrak{N}'$, the desired structure $\mathfrak{A}$ must have, for each $\gamma^K \in$

$\boldsymbol{\gamma}_V^K$, $n(\gamma^K)$ vertices that realize γ^K. Letting $A_{\gamma^K} = \{(i, \gamma^K) \mid 0 \leqslant i < n(\gamma^K)\}$, for all $\gamma^K \in \boldsymbol{\gamma}_V^K$, and setting $V := \bigcup_{\gamma^K \in \boldsymbol{\gamma}_V^K} A_{\gamma^K}$ we therefore put $A := N' \dot{\cup} V$. For all $a \in V$, a will eventually realize in $\mathfrak{A}$ the extended star type of its second component.

The task now is to declare atomic types for all pairs of vertices of A, much as we did in the general case. Let $\widehat{F}$ be the extension of the given F to all of A according to $\widehat{F}(v, (\gamma, \alpha)) := (\gamma, \alpha)$.

First we give a rough sketch of the construction. To give a complete interpretation of a structure over A, all that must be done is to assign atomic 2-types to pairs in $A \setminus K$, i.e., to add $\boldsymbol{\beta}$-edges. This is done in two stages. First, we introduce enough $\boldsymbol{\beta}$-edges such that each $a \in A \setminus K$ has enough outgoing $\boldsymbol{\beta}$-edges to realize the extended star type $\widehat{F}(a)$. (This is the hard part.) In Stage 2, we then fill in the remaining unspecified edges using the (Ramsey theoretic) fact that for each $\gamma_0^K, \gamma_1^K \in \boldsymbol{\gamma}_{\text{lg}}^K$, there is a β such that $\gamma_0^K(\beta) = \gamma_1^K(\overline{\beta}) = 2^+$. (This is easy.)

The basic idea for completing Stage 1 is as follows. (Unfortunately, the 'simple' method does not quite work, and we shall need to use a slightly more complicated construction. But this provides the fundamental intuition.) We enumerate edges in $\boldsymbol{\beta}_{\text{lg}}$ — $\beta_1, \ldots, \beta_n$ in such a way that (i) for all $\beta \in \boldsymbol{\beta}_{\text{lg}}$, there is an $i \leqslant n$ such that $\beta = \beta_i$ or $\overline{\beta} = \beta_i$; (ii) both β and $\overline{\beta}$ do not appear in the list, unless β is symmetric, i.e., $\beta = \overline{\beta}$ (this is because adding β-edges is obviously simultaneously adding $\overline{\beta}$-edges); (iii) for all β_i, β_j, if $\beta_i \in \boldsymbol{\beta}_N$ and $\beta_j \in \boldsymbol{\beta}_V$, then $i < j$, i.e., noble edges appear before villager edges in the list. At the i^{th} step of the process, we want to add enough β_i-edges (connnecting elements 'assigned' by $\widehat{F}$ extended star types in $\Lambda(\beta_i)$ to elements assigned star types in $\Lambda(\overline{\beta}_i)$) so that every such element will have enough outgoing β_i- and $\overline{\beta}_i$-edges to satisfy its assigned extended star type. Also, no a will be given more outgoing β-edges than allowed by the star type $\widehat{F}(a)$.

How do we know that we shall be able to do this in a consistent manner? This is the crux of the entire proof. Suppose that edges have already been added for $\beta_1, \ldots, \beta_{i-1}$, so that, writing β for β_i, we must now add β-edges. To simplify the exposition, we assume that β is separate, so that $V_1 = \{a \in A \mid \widehat{F}(a) = \gamma^K \text{ and } H_\beta(\gamma^K) \neq 0\}$ and $V_2 = \{a \in A \mid \widehat{F}(a) = \gamma^K \text{ and } H_{\overline{\beta}}(\gamma^K) \neq 0\}$ are disjoint. (For example, V_1 is the set of vertices that require an outgoing β-edge.) Let $E = \{(v_1, v_2) \mid v_1 \in V_1, v_2 \in V_2 \text{ and } \mathrm{atp}(v_1, v_2) \text{ has already been determined}\}$. Now, what we have is a bipartite graph, and we need to add β-edges from V_1 to V_2 between pairs that are not already in E, i.e., have not already been assigned an atomic 2-type.

There are 2 cases, depending on whether $\beta \in \boldsymbol{\beta}_N$ or $\in \boldsymbol{\beta}_V$. If $\beta \in \boldsymbol{\beta}_N$, then $V_1, V_2 \subseteq N'$ and we can use the fact that there is a structure $\mathfrak{N}$ that satisfies C4 to determine which β-edges to add. (We have fixed one such $\mathfrak{N}$ throughout the construction.)

On the other hand, if $\beta \in \boldsymbol{\beta}_V$, we know that both V_1 and V_2 are 'big'. Here we use our technical lemmas to show that there is a mapping $f(x)$ from V_1 and V_2 that doesn't overlap with E. Essentially, the β-edges will be the pairs $(v, f(v))$. The reason that we can use the lemmas is that in $A \setminus K$, the *degree* of each $a \in A \setminus K$ — that is, the number of outgoing β_j-edges that it has been assigned — is $\leqslant 2|\boldsymbol{\beta}|$, since this is the maximal number of outgoing edges any element needs to realize its extended star type. Note that we never count edges to kings.

Unfortunately, this is precisely where there is a complication. Consider the case where V_1 is 'too small for' V_2. This can happen precisely when there is a $\gamma^K \in \boldsymbol{\gamma}_{\text{lg}}^K$ such that $\gamma^K(\beta) = 2^+$. For each element in V_2 to have an outgoing $\overline{\beta}$-edge, some $v_1 \in V_1$ must have many outgoing β-edges. But then we cannot use the technical lemmas, since we don't have a uniform bound on the degrees of elements.

Outline of the construction. We now describe how to modify the above plan so that the construction will work. To accomplish the goal of Stage 1, we further divide it into two parts.

Stage 1a: For each $\beta \in \boldsymbol{\beta}_{\text{lg}}$ such that $|V_1|$ is 'too small for' $|V_2|$ (sets V_1, V_2 as described above), we shall choose a distinguished subset $W_\beta \subseteq V_1$, $|W_\beta| = 2|\boldsymbol{\beta}|$ so that for all $w \in W_\beta, \widehat{F}(w) = \gamma^K : \gamma^K(\beta) = 2^+$ and, by the completion of Stage 1a, w will be provided with enough $\boldsymbol{\beta}_{\text{lg}}$-edges to realize the extended star $\widehat{F}(w)$. Also, for all $a \in A \setminus K$ and all $\beta \in \boldsymbol{\beta}_{\text{lg}}$, if a still needs an outgoing β-edge, then there is a set $W_{\overline{\beta}}$ of elements that can accept 2^+ many outgoing $\overline{\beta}$-edges (that is, *incoming* β-edges).

Stage 1b: After Stage 1a, there may still be many $a \in A \setminus K$ that have not been allotted enough outgoing $\boldsymbol{\beta}_{\text{lg}}$-edges. The key idea here is that for each $\beta \in \boldsymbol{\beta}_{\text{lg}}$, all the remaining β-edges can be added so that they connect elements in $A \setminus K$ to elements in $W_{\overline{\beta}}$. But this will only be done after Stage 1a has been completed, so that all the elements in $W_{\overline{\beta}}$ have already been taken care of, i.e., provided with enough outgoing $\boldsymbol{\beta}_{\text{lg}}$-edges. By the construction from Stage 1a, it will be clear that this Stage can be completed easily.

Stage 2: In Stage 1, we added enough outgoing $\boldsymbol{\beta}$-

edges to each $a \in A \setminus K$ for a to realize the star type $\widehat{F}(a)$. But many pairs of elements were not assigned an atomic 2-type, i.e., we did not add an edge connecting them. The construction is completed as in task T3 at the end of the proof of the general case.

The details of the construction are not difficult, but a full account is rather long. In particular, we treat separately a number of different kinds of β-edges in $\boldsymbol{\beta}_{\text{lg}}$. For example, above we discussed a case where β is separate, that is, no villager needs both outgoing β- and $\overline{\beta}$-edges. But we also need to consider cases where β is not separate, where some villagers need both β- and $\overline{\beta}$-edges, while others may need only one or the other. A complete proof will appear in the full paper.

8. C^2-sentences without small models

It was observed above that C^2 does not have the finite model property. We now note that there also exist finitely satisfiable C^2-sentences whose models are necessarily very large, namely double exponential with respect to the length of the sentence.

We construct a family $\{\psi_n : n \in \omega\}$ of C^2-sentences with the following properties:

- ψ_n has a finite model.
- The cardinality of each model of ψ_n is at least 2^{2^n}.
- ψ_n has length $O(n \log n)$ and does not make use of equality.

The vocabulary of ψ_n consists of monadic predicates $P_0, \dots, P_{n-1}, S_0, \dots, S_{n-1}$ and one binary predicate E. The intended model of ψ_n is a binary tree of depth 2^n such that for every node a, the number $s(a) < 2^n$, whose binary representation is given by the sequence of truth values $P_0 a, \dots, P_{n-1} a$, indicates the distance of a from the root. The predicates $S_0, \dots, S_{n-1}$ are auxiliary predicates. The sentence ψ_n is the conjunction of the formulae:

$$\forall x (S_0 x \wedge \bigwedge_{i=1}^{n-1} S_i x \leftrightarrow (S_{i-1} x \wedge P_{i-1} x))$$

$$\exists x \bigwedge_{i=0}^{n-1} \neg P_i x$$

$$\forall x (\bigvee_{i=0}^{n-1} \neg P_i x \rightarrow \exists^{\geqslant 2} y Exy)$$

$$\forall y \exists^{\leqslant 1} x Exy \wedge \forall x \forall y (Exy \rightarrow \beta_n(x,y))$$

where $\beta_n(x,y) := \bigwedge_{i=0}^{n-1} (P_i y \leftrightarrow (P_i x \oplus S_i x))$. Here $\oplus$ stands for the exclusive or. The formula $\beta_n(x,y)$ expresses that $s(y) = s(x) + 1$ modulo 2^n. It is easy to verify that ψ_n has the required properties.

References

[1] L. Babai, P. Erdös, M. Selkov, *Random graph isomorphism*, SIAM Journal of Computing, 9 (1980), pp. 628–635.

[2] L. Babai, L. Kučera, *Canonical labellings of graphs in linear average time*, IEEE Symp. on Foundations of Computer Science (1980), pp. 39-46.

[3] E. Börger, E. Grädel, and Y. Gurevich, *The Classical Decision Problem*, Springer, 1997.

[4] E. Grädel, P. Kolaitis, and M. Vardi, *On the complexity of the decision problem for two-variable first-order logic*, Bulletin of Symbolic Logic, 3 (1997), pp. 53–69. in preparation.

[5] E. Grädel, M. Otto, and E. Rosen, *Undecidability results on two-variable logics*, to appear in Archive of Mathematical Logic, 1999.

[6] R. Graham, B. Rothschild, and J. Spencer, *Ramsey Theory.* John Wiley and Sons, 1980.

[7] N. Immerman and E. Lander, *Describing graphs: a first-order approach to graph canonization*, in A. Selman, ed., Complexity Theory Retrospective, Springer, 1990, pp. 59–81.

[8] M. Mortimer, *On languages with two variables*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 21 (1975), pp. 135–140.

[9] M. Otto, *Bounded variable logics and counting — a study in finite models.* Habilitationsschrift RWTH Aachen, 1995.

[10] ——, *Ptime canonization for two variables with counting*, Proceedings 10th IEEE Symposium on Logic in Computer Science, LICS95, San Diego (1995), 342–352.

[11] L. Pacholski, W. Szwast, and L. Tendera, *Complexity of two-variable logic with counting*, Proceedings 12th IEEE Symposium on Logic in Computer Science, LICS97, Warsaw (1997), 318–327.