

An optimal construction of Hanf formulas¹

Benedikt Bollig² Dietrich Kuske³

²Laboratoire Spécification et Vérification

³Technische Universität Ilmenau

¹Appeared in: *Journal of Applied Logic* **10**:179-186, 2012.

The problem

The construction

Optimality of the construction

Terminology

1. L is a finite relational signature
 L_n is its extension with n constants
2. a d -sphere is an L_n -structure where every element has distance $\leq d$ to some constant, it is **one-centered** if $n = 1$
3. a **sphere** is a d -sphere for some $d \in \mathbb{N}$

Theorem [Hanf '65]

$\forall \mathcal{A}, \mathcal{B}$ locally finite L -structures

$\mathcal{A} \equiv \mathcal{B}$ whenever

any one-centered sphere is realized in \mathcal{A} and in \mathcal{B} the same number of times or infinitely often. (*)

Proof: (*) implies winning strategy for duplicator which implies $\mathcal{A} \equiv \mathcal{B}$ □

Theorem [Fagin, Stockmeyer, Vardi '95]

$\forall r, f \in \mathbb{N} \exists d, m \in \mathbb{N} \forall \mathcal{A}, \mathcal{B}$ L -structures of degree $\leq f$:

$\mathcal{A} \equiv_r \mathcal{B}$ whenever

any one-centered d -sphere is realized in \mathcal{A} and in \mathcal{B}
the same number of times or $\geq m$ times. (*)

Proof: (*) implies winning strategy for duplicator which implies
 $\mathcal{A} \equiv_r \mathcal{B}$ □

Terminology

τ_1, \dots, τ_n all one-centered d -spheres of degree $\leq f$.

For an L -structure \mathcal{A} of degree $\leq f$ and $1 \leq i \leq n$:

$t_i^{\mathcal{A}}$:= minimum of m and number of realisations of τ_i in \mathcal{A} .

$t^{\mathcal{A}} = (t_1^{\mathcal{A}}, t_2^{\mathcal{A}}, \dots, t_n^{\mathcal{A}})$

Corollary

For every sentence φ of quantifier rank $\leq r$, there exist

$T_\varphi, T_\varphi^{\text{fin}} \subseteq \{0, 1, \dots, m\}^n$ such that for all (finite) L -structures \mathcal{A} of degree $\leq f$

$\mathcal{A} \models \varphi$ if and only if $t^{\mathcal{A}} \in T_\varphi$ ($\in T_\varphi^{\text{fin}}$, resp.).

Proof: Collect in T_φ all tuples $t^{\mathcal{A}}$ for (finite) L -structures \mathcal{A} of degree $\leq f$ with $\mathcal{A} \models \varphi$. □

Note: The set T_φ^{fin} from proof is empty iff φ is contradictory. Hence above T_φ^{fin} cannot be computed from φ (Willard '94).

Definition

1. For a d -sphere τ with n centers, the sentence $\text{sph}_\tau(\bar{x})$ expresses “the d -sphere around \bar{x} is isomorphic to τ ”.
2. A formula $\psi(\bar{x})$ is **in Hanf normal form** if it is a Boolean combination of formulas $\exists^{\geq t} x: \text{sph}_\tau(\bar{x}, x)$.
3. Two formulas are **f -equivalent** if they are equivalent on all structures of degree $\leq f$.

Corollary

For every formula φ and every $f \in \mathbb{N}$, there exists an f -equivalent formula in Hanf normal form of triply exponential size.

Corollary [Seese '96]

From a formula φ and $f \in \mathbb{N}$, one can compute an f -equivalent formula in Hanf normal form of triply exponential size.

Proof: Let β express “the structure has degree $\leq f$ ”. Enumerate all tautologies until you find one of the form

$$\beta \rightarrow (\varphi \leftrightarrow \psi)$$

where ψ is in Hanf normal form of triply exponential size. Output ψ . □

Remark

- this construction is not primitive recursive
- Durand & Grandjean '07 and Lindell '08: primitive recursive constructions by elimination of quantifiers and change of signature (superfluous analysis: non-elementary)
- Clochard '12: construction by Durand & Grandjean is 4-fold exponential (3-fold for fixed structure)

The problem

The construction

Optimality of the construction

Let $\varphi(\bar{x}) = Q_1 y_1 Q_2 y_2 \dots \overbrace{Q_{i+1} y_{i+1} \dots Q_n y_n}^{=\varphi_i} : \varphi_n(\bar{x}, y_1, \dots, y_n)$ be a sentence in prenex normal form.

We want to construct inductively and “fast” f -equivalent formulas $\psi_i(\bar{x}, y_1, \dots, y_i)$ in Hanf normal form.

Base step: Let T be the set of all 1-spheres τ of degree $\leq f$ with $|\bar{x}| + n + 1$ centers $\bar{c}, d_1, \dots, d_{n+1}$ s.t.

$$\tau \models \varphi_n(\bar{c}, d_1, \dots, d_n).$$

Then φ_n and $\psi_n = \bigvee_{\tau \in T} \exists^{\geq 1} x_{n+1} : \text{sph}_\tau(\bar{c}, d_1, \dots, d_{n+1})$ are f -equivalent.

Inductive step: $\psi_{i+1}(\bar{x}, y_1, \dots, y_{i+1})$ formula in Hanf normal form (whose radiuses are $\leq d$). We have to transform $\exists y_{i+1}: \psi_{i+1}$ into Hanf normal form.

Then set

$$\psi_i = \bigvee_{\tau} (\alpha_{\tau} \wedge \exists^{\geq 1} y_{i+1} : \text{sph}_{\tau})$$

where the disjunction extends over all $3d$ -spheres of degree $\leq f$ with $|\bar{x}| + i + 1$ centers and α_{τ} is obtained from ψ_{i+1} by replacing

$$\exists^{\geq t} y_{i+2} : \text{sph}_{\tau'}(\bar{x}, y_1, \dots, y_i, y_{i+1}, y_{i+2})$$

by

$$\text{true/false} / \exists^{\geq s} y_{i+2} : \text{sph}_{\sigma}(\bar{x}, y_1, \dots, y_i, \cancel{y_{i+1}}, y_{i+2})$$

where

$$\sigma = \text{Sphere}_{d(\tau')}^{\tau'}(\bar{c}, d_1, \dots, d_i, \cancel{d_{i+1}}, d_{i+2})$$

(analogous to construction of winning strategy). □

Theorem

From a formula φ and $f \geq 1$, one can construct an f -equivalent formula ψ in Hanf normal form of triply exponential size. This construction can be carried out in time

$$2^{f^{2^{O(|\varphi|)}}}.$$

– and Benedikt and I cannot do any better

The problem

The construction

Optimality of the construction

Theorem

There is a family of sentences $(\chi_n)_{n \in \mathbb{N}}$ such that $|\chi_n| \in O(n)$ and every 3-equivalent formula ψ_n in Hanf normal form has $\geq 2^{2^{2^n+1}-1}$ subformulas of the form $\exists^{\geq t} x: \text{sph}_\sigma(x)$, so $|\psi_n| \geq 2^{2^{2^n+1}-1}$.

Proof:

Uses forests \mathcal{A} consisting of binary ordered trees with a unary predicate.

Frick & Grohe '04: there are sentences χ_n of size $O(n)$ s.t. $\mathcal{A} \models \chi_n$ if and only if

no two complete binary trees of height 2^n in \mathcal{A} are isomorphic.

Let ψ_n be in Hanf normal form, 3-equivalent to χ_n , and “small” such that there is one complete binary tree \mathcal{B} with root r and of height 2^n s.t. $\text{sph}_{(\mathcal{B}, r)}$ does not occur in ψ_n .

Let $M \in \mathbb{N}$ be maximal s.t. $\exists^{\geq M}$ appears in ψ_n .

The forest \mathcal{A}_0 consists of $M + 1$ copies of all $\text{Sphere}_d^{\mathcal{B}}(b)$ with $1 \leq d \leq 2^n$ except (\mathcal{B}, r) .

Then \mathcal{A}_0 does not contain any complete binary tree of height 2^n , so $\mathcal{A}_0 \models \psi_n$.

Define $\mathcal{A}_2 = \mathcal{A}_0 \uplus \mathcal{B} \uplus \mathcal{B}$. Then $\mathcal{A}_2 \models \neg\psi_n$, but it realises the same spheres that are mentioned in ψ_n as \mathcal{A}_0 , a contradiction. \square

Summary

Theorem

From a formula φ and $f \geq 1$, one can construct an f -equivalent formula ψ in Hanf normal form of triply exponential size in triply exponential time – and nobody can do any better.