II: Elements of Finite Model Theory

- the differences w.r.t. classical model theory, ranging from 'failures' to new entirely approaches
- more on Ehrenfeucht–Fraïssé: locality techniques for FO, both, classical and with emphasis on FMT
- logic, algorithms and complexity: descriptive complexity & algorithmic model theory
- examples of other logics for specific purposes

just finite structures

shift to finite structures as semantic frame of reference changes the meaning of most semantic notions of logic

σ-structures $\rightsquigarrow$ finite σ-structures, $\mathrm{Fin}(\sigma)$

$$\mathrm{Mod}(\varphi) \rightsquigarrow \mathrm{FMod}(\varphi)$$

$$\varphi \models \psi \rightsquigarrow \varphi \models_{\mathrm{fin}} \psi$$
$$\varphi \equiv \psi \rightsquigarrow \varphi \equiv_{\mathrm{fin}} \psi$$

$$\mathrm{SAT} \rightsquigarrow \mathrm{FINSAT}$$
$$\mathrm{VAL} \rightsquigarrow \mathrm{FINVAL}$$

example: while $\varphi \models \psi$ implies that $\varphi \models_{\mathrm{fin}} \psi$,
converse may fail: "f injective" $\models_{\mathrm{fin}}$ "f surjective"

major differences: compactness and proof systems

no compactness: the FMT analogue of compactness fails for FO

e.g., consider $\Phi := \{\varphi^{\geqslant n} : n \geqslant 1\}$, where $\varphi^{\geqslant n} := \exists x_1 \ldots \exists x_n \bigwedge_{i<j} x_i \neq x_j$

corollary: there cannot be a sound and complete finitistic proof calculus for first-order reasoning about finite models

more specifically compare the cross-over in undecidability results:

Trakhtenbrot Theorem

$\mathrm{FINSAT(FO)}$ and $\mathrm{FINVAL(FO)}$ undecidable;
hence $\mathrm{FINVAL(FO)}$ not r.e., as $\mathrm{FINSAT(FO)}$ is r.e.

Church–Turing Theorem

$\mathrm{SAT(FO)}$ and $\mathrm{VAL(FO)}$ undecidable;
hence $\mathrm{SAT(FO)}$ not r.e., as $\mathrm{VAL(FO)}$ is r.e. (Gödel)

FMT: motivation

finiteness matters:

- sound modelling may require restriction to finite models

  e.g., relational databases correspond to *finite* relational structures

- some issues and phenomena only arise for finite structures

  e.g., asymptotic probabilities; algorithmic and complexity issues concerning structures

variation matters:

- variations in logic and the class of structures go hand–in–hand comparative investigations highlight new issues and yields new methods and new insights into bigger picture

  e.g., combinatorial and algorithmic constructions instead of (inconstructive) compactness arguments; expressive power of many important logics other than FO

examples: 'failures' of classical theorems

substructure preservation

the Łos–Tarski theorem for $\forall$-FO fails in the sense of FMT: there is a FO-sentence whose truth is preserved in the passage to substructures of its finite models, but which is not equivalent in the sense of $\equiv_{\text{fin}}$ (!) to any universal FO-sentence

interpolation

Craig's Interpolation theorem fails for FO in the sense of FMT: there are $\varphi_i \in \mathrm{FO}(\sigma_i)$ s.t. $\varphi_1 \equiv_{\text{fin}} \varphi_2$ that do not admit any interpolant $\chi \in \mathrm{FO}(\sigma_1 \cap \sigma_2)$ s.t. $\varphi_1 \equiv_{\text{fin}} \chi$ and $\chi \equiv_{\text{fin}} \varphi_2$

invariant definability

order invariant FO is more expressive than FO is the sense of FMT: there is $\varphi = \varphi(<) \in \mathrm{FO}_0(\sigma \dot{\cup} \{<\})$ s.t. φ is order-invariant over all finite σ-structures, but not equivalent in the sense of $\equiv_{\text{fin}}$ (!) to any $\mathrm{FO}(\sigma)$-sentence

examples: beyond 'failures'

alternative proofs may yield same/alternative/better results

van Benthem–Rosen

classically and in the sense of FMT, $\mathrm{FO}/\!\sim\, \equiv \mathrm{ML}$:
$\varphi(x) \in \mathrm{FO}$ is preserved under bisimulation, if, and only if, it is equivalent to some $\varphi' \in \mathrm{ML}$
(and one new proof yields same tight bound on quantifier rank in both)

Atserias–Dawar–Grohe

over certain classes of finite structures (e.g., 'wide classes'), the FMT analogue of the Łos–Tarski theorem for $\forall$-FO holds true

Rossman

analogue of classical Lyndon–Tarski theorem for positive $\exists$-FO_{pos} holds true in the sense of FMT; new proof yields classical version plus preservation (!) of quantifier rank (just for classical reading)

II.1 Locality properties of FO

we treat two theorems that play an important role in (non-)expressibility results for FO, especially over finite structure where compactness is not available

- **Gaifman's Theorem** is a theorem of classical model theory: the semantics of relational FO is essentially local
- **Hanf's Theorem** gives a sufficient criterion for $\simeq_m$, over finite structures, in terms of counts of isomorphism types of local neighbourhoods

both can be related to the Ehrenfeucht–Fraïssé analysis of (finite) relational structures in terms of b&f systems based on local conditions

proviso: all signatures finite and relational

locality: Gaifman graph and distance

with a relational σ-structure $\mathfrak{A} = (A, (R^{\mathfrak{A}})_{R\in\sigma})$ associate

its *Gaifman graph* $G(\mathfrak{A}) = (A, E)$ with an E-edge between $a \neq a'$ if a and a' occur together in some tuple $\mathbf{a} \in R^{\mathfrak{A}}$ for some $R \in \sigma$

the *Gaifman distance* on the universe A of $\mathfrak{A}$ as the graph distance $d(a, a') \in \mathbb{N}\dot{\cup}\{\infty\}$ in $G(\mathfrak{A})$

the *Gaifman neighbourhoods* $N^\ell(a)$ of elements $a \in A$ as the subsets $N^\ell(a) = \{a' \in A \colon d(a, a') \leqslant \ell\} \subseteq A$, or also the induced substructures $\mathfrak{A}{\restriction}N^\ell(a)$

for tuples $\mathbf{a} = (a_1, \ldots, a_k)$, put $N^\ell(\mathbf{a}) := \bigcup_{i=1}^k N^\ell(a_i)$

NB: for *finite* σ and fixed $\ell \in \mathbb{N}$,
$d(x, y) \leqslant \ell$, $y \in N^\ell(x)$, $d(x, y) = \ell$, $d(x, y) > \ell$, ...
are expressible by single formulae of $\mathrm{FO}_2(\sigma)$

Hanf and Gaifman theorems

idea:

find sufficient criteria for degrees of FO-equivalence in terms of

- local conditions on equivalences between $\mathfrak{A}\restriction N^\ell(\mathbf{a}), \mathbf{a}$ and $\mathfrak{B}\restriction N^\ell(\mathbf{b}), \mathbf{b}$
- global conditions on $\mathfrak{A}$ and $\mathfrak{B}$

Hanf:

local condition: isomorphism of ℓ_k-neighbourhoods
+ global agreement w.r.t. multiplicities of neighbourhood types

Gaifman:

local condition: $\equiv_{q_k}$-equivalence of ℓ_k-neighbourhoods
+ global agreement w.r.t. scattered tuples for local properties

Hanf's theorem

the ℓ-*neighbourhood type* of an element a in a σ-structure $\mathfrak{A}$ is the isomorphism type of the structure $(\mathfrak{A}\restriction N^\ell(a), a)$;

finite σ-structures $\mathfrak{A}$ and $\mathfrak{B}$ are ℓ-*Hanf-equivalent* if, for every ℓ-neighbourhood type ι,

$|\{a \in A\colon (\mathfrak{A}\restriction N^\ell(a), a) \simeq \iota\}| = |\{b \in B\colon (\mathfrak{B}\restriction N^\ell(b), b) \simeq \iota\}|$

Hanf's theorem

let $\mathfrak{A}, \mathfrak{B}$ be finite σ-structures, σ finite relational; if $\mathfrak{A}$ and $\mathfrak{B}$ are ℓ-Hanf-equivalent for $\ell = \frac{1}{2}(3^{m-1}-1)$, then $(I_k)_{k\leqslant m}\colon \mathfrak{A} \simeq_m \mathfrak{B}$

where $I_m := \{\emptyset\}$ and, for $\ell_k = \frac{1}{2}(3^k-1)$,
$I_k := \big\{p = \mathbf{a} \mapsto \mathbf{b}\colon \mathfrak{A}\restriction N^{\ell_k}(\mathbf{a}), \mathbf{a} \simeq \mathfrak{B}\restriction N^{\ell_k}(\mathbf{b}), \mathbf{b}\big\}$ for $k < m$

typical application:

connectivity of finite graphs not definable in FO nor in $\exists$-MSO

Gaifman's theorem

$\varphi(\mathbf{x}) \in \mathrm{FO}(\sigma)$ is an *ℓ-local formula* if
$\varphi(\mathbf{x}) \equiv \varphi^\ell(\mathbf{x}) := [\varphi(\mathbf{x})]^{N^\ell(\mathbf{x})}$ (relativisation to $N^\ell(\mathbf{x})$),
i.e., for all $\mathfrak{A}, \mathbf{a}$: $\mathfrak{A}, \mathbf{a} \models \varphi$ iff $\mathfrak{A} \restriction N^\ell(\mathbf{a}), \mathbf{a} \models \varphi$

a *basic ℓ-local sentence* is an FO-sentence of the form

$$\varphi = \exists x_1 \ldots \exists x_m \big(\textstyle\bigwedge_{i<j} d(x_i, x_j) > 2\ell \;\wedge\; \bigwedge_i \psi^\ell(x_i)\big)$$

for some ℓ-local formula $\psi^\ell(x) \in \mathrm{FO}_1(\sigma)$

NB: the following is a theorem of classical model theory

Gaifman's theorem

for any relational signature σ, every $\varphi(\mathbf{x}) \in \mathrm{FO}(\sigma)$
is logically equivalent to a boolean combination of
local formulae and basic local sentences

for a b&f proof of Gaifman's theorem:

the rank of the basic local sentence
$\varphi = \exists x_1 \ldots \exists x_m \big(\bigwedge_{i<j} d(x_i, x_j) > 2\ell \;\wedge\; \bigwedge_i \psi^\ell(x_i)\big)$ is $(\ell, \mathrm{qr}(\psi), m)$

definition

$\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ are (ℓ, q, m)-*Gaifman-equivalent*, $\mathfrak{A}, \mathbf{a} \equiv^\ell_{q,m} \mathfrak{B}, \mathbf{b}$, if

- $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ satisfy the same ℓ-local formulae of quantifier rank up to q;
- $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same basic local sentences in ranks (ℓ', q', m') for $\ell \leqslant \ell$, $q' \leqslant q$, $m' \leqslant m$

lemma

if $\mathfrak{A}$ and $\mathfrak{B}$ are $(L, Q, m+n)$-Gaifman-equivalent for sufficiently large L, Q, $\mathbf{a}_0 \in A^n, \mathbf{b}_0 \in B^n$, then $(I_k)_{k \leqslant m} \colon \mathfrak{A}, \mathbf{a}_0 \simeq_m \mathfrak{B}, \mathbf{b}_0$

where, for suitable (ℓ_k, q_k), I_k consists of all partial isomorphisms $p = \mathbf{a} \mapsto \mathbf{b}$, $|p| \leqslant m + n - k$, s.t. $\mathfrak{A} \restriction N^{\ell_k}(\mathbf{a}), \mathbf{a} \equiv_{q_k} \mathfrak{B} \restriction N^{\ell_k}(\mathbf{b}), \mathbf{b}$