II.3 FO^k: the k-variable fragment of FO

in relational signature σ, formulae of $\mathrm{FO}^k(\sigma) \subseteq \mathrm{FO}(\sigma)$
use only k distinct variable symbols $(x_1, \ldots, x_k)$ throughout,
re-usable in nested quantifications, as in

$\varphi(x) := \exists y\big(Exy \wedge \exists x\big(Eyx \wedge \exists y\big(Exy \wedge \exists x Eyx\big)\big)\big) \in \mathrm{FO}^2(E)$,
which says that there is an E-path of length 4 from x

NB: subformulae of an FO^k-formula define and can be evaluated
in terms of relations of arity up to k only

Remark: FO^2 has the finite model property, whence
$\mathrm{SAT}(\mathrm{FO}^2) = \mathrm{FINSAT}(\mathrm{FO}^2)$ is decidable (Mortimer);
in fact, FO^2 even satisfies a small model property with
an exponential bound on the size of minimal size models
and $\mathrm{SAT}(\mathrm{FO}^2)$ is in NExptime (Grädel–Kolaitis–Vardi)

FO^k and k-pebble games

adaptation of FO-Ehrenfeucht–Fraïssé game and b&f notions:
configurations $(\mathfrak{A}, \mathbf{a}; \mathfrak{B}, \mathbf{b})$ with $\mathbf{a} \in A^k, \mathbf{b} \in B^k$

in single round:

- player **I** selects pebble pair to be relocated
  and moves corresponding pebble in one structure
- player **II** must match move in opposite structure
  and maintain partial isomorphisms of size (up to) k

notions of k-pebble game equivalence, $\simeq^k_m$ and $\simeq^k_\infty$ ($=\simeq^k_{\mathrm{part}}$), and
corresponding b&f systems with suitable k-pebble b&f conditions,
are naturally defined; a k-pebble Ehrenfeucht–Fraïssé theorem is
obtained for structures in finite relational signature

inductve refinement process yields minimal m such that $\simeq_{m+1}$
coincides with $\simeq_m$ (and thus with $\simeq^k_\infty$) on $\mathfrak{A}$: the k-rank of $\mathfrak{A}$,
which is bounded by $|A|^k$

complete k-pebble invariants

inductive pre-order refinement of levels $\simeq^k_m$ over individual $\mathfrak{A}$
yields a linearly ordered invariant $\mathfrak{I}^k(\mathfrak{A}, \mathbf{a})$
as a structural abstraction of $(\mathfrak{A}, \mathbf{a})/ \simeq^k_\infty$

based on inductive refinement process
of levels $A^k/(\simeq^k_m)^{\mathfrak{A}} \longrightarrow A^k/(\simeq^k_\infty)^{\mathfrak{A}}$,
sorting in new classes lexicographically,
terminating within k-rank of $\mathfrak{A}$ many steps

for fixed finite relational σ, $\mathfrak{I}^k$ provides concise, Ptime computable complete invariant w.r.t. $\simeq^k_\infty$ over $\mathrm{Fin}(\sigma)$:

$$\boxed{\begin{array}{l}\text{for all } \mathfrak{A}, \mathfrak{B} \in \mathrm{Fin}(\sigma), \mathbf{a} \in A^k, \mathbf{b} \in B^k:\\ \mathfrak{A}, \mathbf{a} \simeq^k_\infty \mathfrak{B}, \mathbf{b} \;\Leftrightarrow\; \mathfrak{I}^k(\mathfrak{A}, \mathbf{a}) \simeq \mathfrak{I}^k(\mathfrak{B}, \mathbf{b})\end{array}} \quad (*)$$

due to its linearly ordered nature, it is 'essentially syntactic',
i.e., we could replace $\simeq$ by $=$ in $(*)$ after normalisation

II.4 Fixpoint logics in descriptive complexity

(A) descriptive comlexity:

logical (or other machine-independent)
characterisations of complexity classes

towards an alternative analysis and understanding
of the levels of algorithmic complexity of problems

example: Büchi's theorem, giving a precise match

computational power of finite automata	$\leftrightsquigarrow$	expressive power of MSO

FMT is concerned with the complexity of structural problems,
especially decision problems based on properties of structures;

the study of these (boolean) *queries* is *richer* than the standard setting, since coding & representation impose semantic constraints

review: computational complexity

NB: complexity classes are classes of problems, not of algorithms;
defined in terms of resource bounds on Turing machines
(think: "worst-case complexity of best possible machine")

(B) standard complexity classes:

P/Ptime: polynomial time, termination within $p(n)$ steps on inputs of size n, for some polynomial p

NP: non-deterministic polynomial time, based on polynomial depth non-deterministic procedure

potentially exponentially branching search & verification,
or: guessing of polynomial size certificate and Ptime check

Pspace: polynomial space, termination with polynomially bounded overall memory (tape) consumption

potentially exponential time

the need for coding and the role of order

(C) structures as inputs; queries

standard input for Turing machines (or for standard algorithms) are strings/words over some suitable finite alphabet

input structures $\mathfrak{A} \in \mathrm{Fin}(\sigma)$ have to be encoded as words;
as part of the correctness condition on admissible algorithms,
different encodings of the same (or isomorphic) structures
as input have to lead to the same output result:

- queries on $\mathrm{Fin}(\sigma)$ are, by definition, $\simeq$-invariant;
  this is a non-trivial semantic constraint, which
  is computationally non-trivial below NP
- linearly ordered structures admit canonical encoding schemes
  that are unambiguous, thus trivialising the issue

  $\rightsquigarrow$ the crucial role of order-invariance in FMT

Fagin's theorem

observation: FO-definable queries are in Ptime (even Logspace)

theorem (Fagin)

NP $\equiv$ $\exists$-**SO**, existential second-order logic captures NP,
i.e., the following are equivalent for all $\mathcal{C} \subseteq \mathrm{Fin}(\sigma)$:

(i) $\mathcal{C}$ is definable by a sentence of existential second-order logic:
$\mathcal{C} = \mathrm{FMod}(\exists \mathbf{X} \varphi(\mathbf{X}))$ for some $\varphi(\mathbf{X}) \in \mathrm{FO}(\sigma \cup \{\mathbf{X}\})$

(ii) the decision problem for $\mathcal{C} \subseteq \mathrm{Fin}(\sigma)$ is in NP

NB: this is an assertion on the model checking complexity of $\exists$-SO, together with a matching expressive completeness result for $\exists$-SO !

NB: order is dispensable, since available in existential quantification

coding of configurations and runs

encode n^k-bounded runs of (non-deterministic) TM
$\mathcal{M} = (\Gamma, Q, q_0, q^+, q^-, \Delta)$ on input structures
$(\mathfrak{A}, <) = (\{0, \ldots, |A| - 1\}, <, \ldots)$ with linear ordering $<$
over $(n, <) = (\{0, \ldots, n-1\}, <)$ for $n = |A|$

- use A^k as numerical domain for numbers encoded to base n
- encode run $(C_t)_{t<n^k}$ with $C_t = (q_t, \ell_t, \rho_t)$
  as the graph of a function $A^k \times A^k \to A \times A^k \times A$
  i.e., as a relation $R \subseteq A^{3k+2}$

  initial configuration C_0
  consistency of $C_t \rightsquigarrow C_{t+1}$ with Δ
  accepting final state
  } in $\mathrm{FO}(\sigma \cup \{R, <\})$

Fagin: (implicit) FO-definability $\rightsquigarrow$ explicit $\exists$-SO definability
also without (the invariant use of) order

other capturing results with order $\to$ below

relational recursion: fixpoint logics

$\varphi(X, \mathbf{x}) \in \mathrm{FO}_k(\sigma \cup \{X\})$ with k-ary X and matching $\mathbf{x}$ induces operation on $\mathcal{P}(A^k)$, uniformly across all $\mathfrak{A} \in \mathrm{Fin}(\sigma)$:

$$\begin{aligned} \mathcal{F}^{\mathfrak{A}}_{\varphi} \colon \mathcal{P}(A^k) &\longrightarrow \mathcal{P}(A^k) \\ P &\longmapsto \{\mathbf{a} \in A^k : \mathfrak{A}, P, \mathbf{a} \models \varphi\} \end{aligned}$$

easy to see: if $\varphi(X, \mathbf{x})$ is X-positive, this operation is monotone (preservation result/classically only: matching expressive completeness)

natural extensions of FO, esp. for FMT, provide recursion mechanisms based on such definable operations

- **least fixpoint logic LFP** has least and greatest fixpoints for positive/monotone operations
- **partial fixpoint logic PFP** has fixpoints for arbitrary operations (with default $\emptyset$)

capturing results with order

thm (Immerman–Vardi)

Ptime $\equiv$ LFP over linearly ordered structures

i.e., t.f.a.e. for every class $\mathcal{C} \subseteq \mathrm{Fin}(\sigma)$ of linearly ordered σ-structures:

(i) $\mathcal{C} \subseteq \mathrm{Fin}(\sigma)$ is decidable in NP

(ii) $\mathcal{C}$ is definable within $\mathrm{Fin}(\sigma)$ by a sentence of $\mathrm{LFP}(\sigma)$

thm (Abiteboul–Vianu)

Pspace $\equiv$ PFP over linearly ordered structures

remarks: order is crucial, simple fixpoints over FO suffice
model-checking in Ptime/Pspace is obvious
expressive completeness: coding & fixpoint recursion