FIXED-POINT EXTENSIONS OF FIRST-ORDER LOGIC

Yuri Gurevich* and Saharon Shelah**

Electrical Engineering and Computer Science Dept.
University of Michigan, Ann Arbor, MI 48109-1109

Institute of Mathematics and Computer Science
Hebrew University, 91904 Jerusalem, Israel, and
EECS and Mathematics Departments
University of Michigan, Ann Arbor, Michigan

Abstract. We prove that the three extensions of first-order logic by means of positive inductions, monotone inductions, and so-called non-monotone (in our terminology, inflationary) inductions respectively, all have the same expressive power in the case of finite structures. As a by-product, the collapse of the corresponding fixed-point hierarchies can be deduced.

§0. Introduction

In 1979 Aho and Ullman [AU] noted that the relational calculus is unable to express the transitive closure, and suggested extending the relational calculus by the least fixed point construct. The relational calculus [Ul] is a standard relational query language; from the point of view of expressive power, the relational calculus is exactly first-order logic. Aho and Ullman's paper triggered an extensive study of the expressive power of fixed-point extensions of first-order logic [CH, Im1, Va, Li, Gu, BGK, etc.] with emphasis on finite structures.

There are two fields where fixed-point extension of first-order logic were extensively studied earlier. One is the theory of inductive definitions summarized to an extent in the book [Mo]. The other is semantics of programming languages where a fixed-point extension of first-order logic is known as first-order μ-calculus. But neither of the two fields put finite structures into the center of attention.

* Partially supported by NSF grant MCS 83-01022

** Partially supported by NSF grant MCS 81-01560

Proviso. All structures are finite unless the contrary is said explicitly.

Fixed-point constructions arise in the frame of first-order logic quite naturally. A formula $\varphi(P,x)$ with an r-ary predicate variable P and a sequence x of r free individual variables yields an operator $F(P)=\{x: \varphi(P,x)\}$ that can be applied repetitively. Additional free variables of φ are viewed as parameters. If F is monotone then it has a least (with respect to the inclusion relation) fixed point $LFP(F)=LFP_{P;x}\varphi(P,x)=\cup F^i(\emptyset)$.

E.g., $LFP_{P;x,y}(Edge(x,y) \text{ or } \exists z[P(x,z) \,\&\, P(z,y)])$ is the transitive closure of Edge, and $LFP_{P;x}(x=u \text{ or } x=v \text{ or } \exists y \exists z[P(y) \,\&\, P(z) \,\&\, x=f(y,z)])$ is the closure of set $\{u,v\}$ under the operation f.

Unfortunately, the extension of first-order logic by the construct LFP applicable to formulas φ with a monotone F, is not a nice logic because recognizing well-formed formulas is undecidable [Gu]. But there is a simply recognizable sufficient condition for monotonicity. If a first-order $\varphi(P,x)$ is positive in P then the operator $F(P)=\{x: \varphi(P,x)\}$ is monotone. Moreover, the definition of positivity naturally extends to new formulas, and positivity remains sufficient for monotonicity. The extension FO+LFP of first-order logic by the constuct LFP, applicable to positive formulas, is most popular.

The restriction to positive formulas has its own price. In many cases it is obvious that a given formula $\varphi(P,x)$ yields a monotone operator F but it is not clear how to transform $\varphi(P,x)$ to an equivalent formula $\varphi'(P,x)$ which is positive in P. (A first-order $\varphi(P)$ may yield a monotone operator and have no first-order equivalent $\varphi'(P)$ that is positive in P [AG].) In order to define a more flexible fixed-point extension of first-order logic, it is worth to loosen the condition of monotonicity rather than to tighten it up.

Call an operator $F(P)=\{x: \varphi(P,x)\}$ <u>inductive</u> if the sequence $F^i(\emptyset)$ increases. If F is inductive then $\cup F^i(\emptyset)$ is a (not necessarily the least) fixed point of F that will be called the <u>inductive fixed point</u> $IFP(F)=IFP_{P;x}\varphi(P,x)$ of F. Call F <u>inflationary</u> if $\forall P[P \subseteq F(P)]$. Any inflationary F is inductive. The operator $F'(P)=\{x: P(x) \text{ or } \varphi(P,x)\}$ is inflationary, and if F is monotone then $IFP(F')=LFP(F)$. This suggests an extension FO+IFP [Gu, Li] of first-order logic by the construct IFP applicable to any formula [P(x) or $\varphi(P,x)$] with arity(P)=length(x).

Obviously, FO+LFP ≤ FO+IFP by expressive power, and the monotonicity bound extension lies in-between. Every FO+IFP query is computable within time polynomial in the size of a given structure. In the presence of linear order, every polynomial time computable relational query is expressible in FO+LFP [Im, Va]; the presence of order allows to simulate Turing machines. Thus, in the presence of linear order, FO+LFP and FO+IFP have the same expressive power. In general, however, not every polynomial time computable query is expressible in FO+LFP [CH] or even FO+IFP [BGK]. The general case is important: a query may depend not on specifics of the given representation but only on the isomorphism type of the given structure. We show that even in the general case FO+LFP and FO+IFP have the same expressive power. Actually, a stronger result holds.

<u>Theorem</u> 1 (Main Theorem). Let Γ be an arbitrary operator that, given two r-ary relations and an r-tuple of elements, produces a boolean value. Then

$$IFP_{P;x}[P(x) \text{ or } \Gamma(P,-P,x)] = LFP_{Q;y}\psi(Q,y)$$

for some ψ which is built from Γ by first-order means and is positive in the predicate variable Q.

It is supposed of course that $\Gamma(P,P',x)$ is positive in both predicate variables. In applications, given a formula $\gamma(P,x)$, define $\Gamma(P,P',x)$ as the result of substituting P' for the negative occurrences of P in γ. Main Theorem speaks about arbitrary $\Gamma(P,-P,x)$ rather than arbitrary $\gamma(P,x)$ because of the need to distinguish between positive and negative occurrences of P. Apart from this, the internal structure of the given formula is of no importance in constructing the desired ψ.

<u>Corollary</u> 1. FO+LFP, FO+IFP and the monotonicity bound extension of first-order logic all have the same expressive power.

<u>Corollary</u> 2. For every first-order formula $\gamma(P,x)$ there is a first-order formula $\psi(Q,y)$ such that $\psi(Q,y)$ is positive in Q and

$$IFP_{P;x}[P(x) \text{ or } \gamma(P,x)] = LFP_{Q;y}\psi(Q,y).$$

The proof of Main Theorem is sketched in §3; the full proof of Main Theorem will appear in [Gu].

Chandra and Harel [CH] raised the question about the LFP hierarchy in logic FO+LFP. Immerman [Im1] announced that the LFP hierarchy collapses on the first level; he elaborated his solution in [Im2]. In July 1985, Phokion Kolaitis brought to our attention some difficulties in Immeman's proof. We saw immediately that the IFP hierarchy collapses on the first level.

<u>Theorem</u> 2 (See §4). Every FO+IFP formula is equivalent to an FO+IFP formula which is either first-order or of the form $[IFP_{...}\Phi](...)$ where Φ is first-order.

Moreover, the proof of hierarchy collapse is very natural in the FO+IFP setting. Immerman told us that all difficulties will be taken care of in a new version of [Im2]. Anyway, Theorem 2 and Corollary 2 imply

<u>Corollary</u> 3 [Cf. Im1, Im2]. Every FO+LFP formula is equivalent to an FO+LFP formula which is either first-order or of the form $[LFP_{...}\Phi](...)$ where Φ is first-order.

Acknowledgement. We thank Egon Boerger, Alekos Kechris, Phokion Kolaitis, Yiannis Moschovakis and Dana Scott for useful discussions.

§1. Defining logics FO+LFP and FO+IFP

In this section structures are not necessarily finite.

Definition. Let **P** be a complete partially ordered set and F be a function from **P** to **P**. Let P_0=min(**P**), $P_{\alpha+1}=F(P_\alpha)$, and $P_\alpha=\sup\{P_\beta: \beta<\alpha\}$ for limit α. If the sequence P_α is increasing (i.e. $\alpha<\beta \to P_\alpha \le P_\beta$) then F is inductive. If F is inductive and $\mu=\min\{\alpha: P_\alpha=P_{\alpha+1}\}$ then P_μ is the inductive fixed point IFP(F) of F. If $X \le F(X)$ for every $X \in$ **P** then F is inflationary.

Theorem 1. Let **P** be a complete partially ordered set and F: **P** → **P**.

(a) If F is inflationary then it is inductive.

(b) The function sup{X,F(X)} is inflationary; its inductive fixed point equals IFP(F) if F is inductive.

(c) If F is monotone (i.e. $X \le Y \to F(X) \le F(Y)$) then F is inductive and the inductive fixed point of F is a least fixed point LFP(F) of F.

Proof is clear. □

Examples. Suppose U={0,1, 2}, **P** is the power set of U ordered by inclusion, and X ranges over **P**.

(i) Let F(X)=X∪{the cardinality of X} if X≠U, and F(U)=U. Then F is inflationary but does not have a least fixed point: {1} and {0,2} are fixed points of F but ∅ is not.

(ii) Let G(X)=F(X) if X is an initial segment of U, and G(X)=∅ otherwise. Then G is inductive but neither inflationary nor monotone.

(iii) A constant function H(X)={0} is monotone but not inflationary.

The syntax of logic FO+LFP is the result of augmenting the syntax of first-order logic by:

LFP Formation Rule. Let r be a positive integer, **x** be an r-tuple $x_1,\dots,x_r$ of individual variables, P be an r-ary predicate variable, $\varphi(P,x)$ be a well-formed formula, and t be an r-tuple of terms. If $\varphi(P,x)$ is positive in P (i.e. all free occurrences of P in $\varphi(P,x)$ are positive) then $LFP_{P;x}\varphi(P,x)$ is a well-formed predicate and $[LFP_{P;x}\varphi(P,x)](t)$ is a well-formed formula.

P and $x_1,\dots,x_r$ are bounded in the new predicate. Other free individual or predicate variables of φ remain free in the new predicate. If Q is a predicate variable different from P then every positive (respectively, negative) occurrence of Q in $\varphi(P,x)$ remains positive (respectively, negative) in the new predicate.

Remark. A simplified notation $LFP_P\varphi(P,x)$ for $[LFP_{P;x}\varphi(P,x)](x)$ is deficient: just try to express $[LFP_{P;x}\varphi(P,x)](t)$ in the simplified notation.

To be on the safe side, let us emphasize that logic FO+LFP allows interleaving LFP with propositional connectives (including negation) and quantifiers; in particular, one can negate an LFP formula then use the LFP formation rule again, etc.

The meaning of the predicate $LFP_{P;x}\varphi(P,x)$ is the least fixed point of the operator $F(P)=\{x: \varphi(P,x)\}$ on the set of r-place predicates ordered by inclusion. Since the formula $\varphi(P,x)$ is positive in P, the operator F is monotone and therefore has a least fixed point.

As we have mentioned in the introduction, direct replacement of positivity by monotonicity in the LFP formation rule does not lead to a nice logic. However, the operator $F'(P)=\{x: P(x) \text{ or } \varphi(P,x)\}$ is always inflationary and therefore has an inductive fixed point. By Theorem 1, IFP(F')=LFP(F) if F is monotone. This leads to a more liberal extension FO+IFP of first-order logic. Let us call a formula $\varphi(P,x)$ (in whatever language) explicitly inflationary if $\varphi(P,x)=[P(x) \text{ or } \Phi(P,x)]$ for some Φ. The syntax of logic FO+IFP is the result of augmenting the syntax of first-order logic by:

IFP Formation Rule. Let r be a positive integer, **x** be an r-tuple of individual variables, P be an r-ary predicate variable, $\varphi(P,x)$ be an arbitrary well-formed formula, and t be an r-tuple of terms.

If the formula $\varphi(P,x)$ is explicitly inflationary then $IFP_{P;x}\varphi(P,x)$ is a well-formed predicate, and $[IFP_{P;x}\varphi(P,x)](t)$ is a well-formed formula.

The meaning of the predicate $IFP_{P;x}\varphi(P,x)$ is the inductive fixed point of the inflationary operator $F(P)=\{x: \varphi(P,x)\}$.

§2. Simultaneous induction

For reader's convenience we prove in this section the known fact that simultaneous induction reduces to the ordinary one. Structures are not necessarily finite.

Given natural numbers p and q, order the set $\{(P,Q)$: P is a p-ary predicate and Q is a q-ary predicate$\}$ componentwise: $(P,Q) \leq (P',Q')$ if $P \subseteq P'$ and $Q \subseteq Q'$. The resulting partially ordered set is complete. Let x and y be sequences of individual variables of length p and q respectively.

Simultaneous Induction Lemma for FO+LFP [Cf. Mo]. Let $F(P,Q) = (\{x: \varphi(P,Q,x)\}, \{y: \psi(P,Q,y\})$ be an operator where φ, ψ are FO+LFP formulas positive in P and Q. Let

$(LFP1_{P,Q;x,y}(\varphi,\psi), LFP2_{P,Q;x,y}(\varphi,\psi))$

be the least fixed point of F. Then there is an FO+LFP formula $\alpha(x)$ such that

$\alpha(x) \leftrightarrow [LFP1_{P,Q;x,y}(\varphi,\psi)](x)$, and

$\alpha(x)$ has the form $[LFP_{...}\gamma](...)$ where γ is built from φ,ψ by first-order means.

Proof. To simplify the exposition we suppose that $x=(x_1,x_2)$ and $y=(y_1,y_2,y_3)$. Let u,v,w,w' be individual variables, R be a new predicate variable of arity 5=2+max{p,q}, and z be a triple (z_1,z_2,z_3) of new individual variables. Let $\gamma(R,u,v,z_1,z_2,z_3)$ say the following:

Either there is only one element in the universe, and an equivalent of $[LFP1_{P,Q;x,y}(\varphi,\psi)](z_1,z_2)$, built from φ and ψ by first-order means, holds,

or there are $w \neq w'$ such that

$u=v=z_3$ & $\varphi(\{x: R(w,w,x,w)\}, \{y: R(w,w',y)\}, z_1,z_2)$

or

$u \neq v$ & $\psi(\{x: R(w,w,x,w)\}, \{y: R(w,w',y)\}, z_1,z_2,z_3)$.

The idea is to represent P(x) by R(u,u,x,u) with arbitrary u, and Q(y) by R(u,v,y) with arbitrary $u \neq v$. The desired $\alpha(x) = [LFP_{R;u,v,z}\gamma](x_1,x_1,x,x_1)$. If p=0 then $\alpha = [LFP_{R;u,v,y}\gamma](x_1,\dots,x_1)$ where x_1 is a new variable or a constant. □

The proof is a slight modification of the corresponding proof in [Mo]. (The possibility of using only one individual constant or even none at all is mentioned in [Im2].) The same proof establishes

Simultaneous Induction Lemma for FO+IFP. Let $F(P,Q) = (\{x: \varphi(P,Q,x)\}, \{y: \psi(P,Q,y)\})$ be an operator where φ, ψ are explicitly inflationary FO+IFP formulas. Let

$(IFP1_{P,Q;x,y}(\varphi,\psi), IFP2_{P,Q;x,y}(\varphi,\psi))$

be the inductive fixed point of F. Then there is an FO+IFP formula $\alpha(x)$ such that

$\alpha(x) \leftrightarrow [IFP1_{P,Q;x,y}(\varphi,\psi)](x)$, and

$\alpha(x)$ has the form $[IFP_{...}\gamma](...)$ where γ is built from φ,ψ by first-order means.

It is easy to formulate and prove analogues of the two lemmas for the case when three or more relations are defined by simultaneous induction. In a sense, fixed-point logics with built-in simultaneous induction are more natural. In the sequel we will use the extension of FO+LFP by an additional formation rule for $LFP1_{P,Q;x,y}(\varphi,\psi)$, and the extension of FO+IFP by additional formation rules for $IFP1_{P,Q;x,y}(\varphi,\psi)$ and $IFP1_{P,Q,R;x,y,z}(\varphi,\psi,\chi)$. By the simultaneous induction lemmas, the additional formation rules do not increase the expressive power.

§3. Expressing the inductive fixed point

The proviso of §0 is in force: all structures are finite.

<u>Theorem</u> 1. Let Γ be an arbitrary operator that, given two unary relations and an element, produces a boolean value. Then

$$[\mathrm{IFP}_{P;x}(P(x) \text{ or } \Gamma(P,-P,x))](x)$$

is equivalent to a formula

$$([\mathrm{LFP}1_{R,S;x,y,z,u,v,w}(\rho,\sigma)](x,x,x))$$

where ρ,σ are built from Γ by first-order means and are positive in the ternary predicate variables R,S.

We write $\Gamma(P,-P,x)$ rather than $\gamma(P,x)$ in order to distinguish between positive and negative occurrences of P. First-order formulas and formulas built from Γ by first-order means will be called <u>pseudo first-order</u>. The notion of positivity is generalized to pseudo first-order formulas in the obvious way; in particular the pseudo first-order formula $\Gamma(P,P',x)$ is positive in both P and P'.

<u>Corollary</u> 2. Theorem 1 remains true under the vector interpretation (when x is interpreted as an r-tuple of individual variables, P is interpreted as an r-ary predicate variable and so on).

<u>Corollary</u> 3. Every FO+IFP formula is equivalent to an appropriate FO+LFP formula.

<u>Proof</u> of Corollary 3 proceeds by induction. The only non-trivial case is that of $[\mathrm{IFP}_{P;x}(P(x)$ or $\gamma(P,x)](x)$ where γ – by the induction hypothesis – can be assumed to be an FO+LFP formula. Let $\Gamma(P,P',x)$ be the result of replacing the negative occurrences of P in $\gamma(P,x)$ by $-P'$ where P' is a new predicate variable. Obviously, $\Gamma(P,-P,x) \leftrightarrow \gamma(P,x)$. Now use Corollary 2. □

In the rest of this section we sketch a proof of Theorem 1. For expositary purposes we choose a nonempty finite set U as our universe of discourse. Let $\varphi(P,x)=[P(x)$ or $\Gamma(P,-P,x)]$, $F(P)=\{x: \varphi(P,x)\}$ and $P_n=F^n(\emptyset)$ i.e. $P_0=\emptyset$ and $P_{n+1}=F(P_n)$. The sequence P_n is (non-strictly) increasing. Let $m=\min\{n: P_n=P_{n+1}\}$; P_m is the inductive fixed point of F. In addition, let $P_\infty=U$. For every $x\in U$, let $\mathrm{stage}(x)=\min\{n: x\in P_n\}$. Note that $\mathrm{stage}(x)>0$. Let $x\le y$ abbreviate $[x\in P_m$ and $\mathrm{stage}(x)\le \mathrm{stage}(y)]$, and let $x<y$ abbreviate $\mathrm{stage}(x)<\mathrm{stage}(y)$. Note that $x\le x \leftrightarrow x\in P_m$. We start with defining an auxiliary inductive operator G whose inductive fixed point is the relation $\le$.

<u>Lemma</u> 4 [Cf. Stage Comparison Theorem, Mo].
$x\le y \leftrightarrow \varphi(\{x': x'<y\},x)$,
$x<y \leftrightarrow -\varphi(\{y': -(x\le y')\}, y)$, and
$x\le y \leftrightarrow \varphi(\{x': -\varphi[\{y': -(x'\le y')\}, y]\}, x)$.

The proof is straightforward; formally speaking, the lemma will not be used. The last statement of Lemma 4 gives the desired G, but the need to keep a track of the positive and negative occurrences of the induction variable forces us to give a more explicit definition of G.

Let Q, Q' be binary predicates variables. Let

$$\Delta(Q,Q',x',y)=[Q'(x',y) \text{ or } \Gamma(\{y': Q'(x',y')\}, \{y': Q(x',y')\}, y)],$$
$$\Delta'(Q,Q',x',y)=-\Delta(-Q',-Q,x',y),$$
$$\Psi(Q,Q',x,y)=[\Delta'(Q,Q',x,y) \text{ or } \Gamma(\{x': \Delta'(Q,Q',x',y)\}, \{x': \Delta(Q,Q',x',y)\}, x)],$$
and $G(Q)=\{(x,y): \Psi(Q,-Q,x,y)\}$.

Then Δ, Δ' and Ψ are positive in Q and Q', and

$$\Delta(Q,-Q,x',y) \leftrightarrow \varphi(\{y': -Q(x',y')\},y).$$

<u>Lemma</u> 5. $\Psi(Q,-Q,x,y) \leftrightarrow \varphi(\{x': -\varphi[\{y': -(x'\le y')\}, y]\}, x)$.

<u>Proof</u>. $[\Delta'(Q,-Q,x,y)$ or
$\Gamma(\{x': \Delta'(Q,-Q,x',y)\}, \{x': \Delta(Q,-Q,x',y)\}, x)] \leftrightarrow$
$[-\Delta(Q,-Q,x,y)$ or
$\Gamma(\{x': -\Delta(Q,-Q,x',y)\}, \{x': \Delta(Q,-Q,x',y)\}, x)] \leftrightarrow$
$\varphi(\{x': -\Delta(Q,-Q,x',y)\},x) \leftrightarrow$
$\varphi(\{x': -\varphi(\{y': -Q(x',y')\},y)\},x)$. □

Let $Q_k=G^k(\emptyset)$. We show that Q_k's are approximations to $\le$.

<u>Lemma</u> 6. For every natural number k, $Q_k=\cup\{(P_i\times P_\beta): k\ge i\le\beta\}$ where β may be equal to ∞.

<u>Proof</u> by induction on k. Case k=0 is clear. We suppose $Q_k=\cup\{(P_i\times P_\beta): k\ge i\le\beta\}$ and prove $Q_{k+1}=\cup\{(P_i\times P_\beta): k+1\ge i\le\beta\}$.

First, check that $-\varphi(\{y': -Q_k(x',y')\}, y)$ holds if and only if stage(y)>stage(x')$\leq$k. Second, let β=stage(y). We have $\{x': -\varphi(\{y': -Q_k(x',y')\}, y)\}=\{x': \beta>\text{stage}(x')\leq k\}=P_j$ where $j+1=\min\{\beta,k+1\}$. Third, let i=stage(x). Then $(x,y)\epsilon Q_{k+1} \leftrightarrow \varphi(P_j,x) \leftrightarrow i\leq j+1 \leftrightarrow (i\leq\beta$ and $i\leq k+1) \leftrightarrow (x,y)\epsilon\cup\{(P_i\times P_\beta): \beta\geq i\leq k+1\}$. □

Corollary 7. The operator G is inductive, Q_m is the inductive fixed point of G, and the relation $\leq$ coincides with Q_m.

Now we come to the crucial transition to a positive induction. Note that the formula $\Psi(Q,-Q,x,y)$ is, in general, not positive in Q; it defines Q_{k+1} in terms of Q_k and $-Q_k$. Our idea is to build, by a positive simultaneous induction, two ternary relations R and S in such a way that

$$R_{k+1}-R_k=\{k+1\} \times Q_{k+1},$$
$$S_{k+1}-S_k=\{k+1\}\times -Q_{k+1}.$$

This would allow us to use positive occurrences of S instead of negative occurrences of Q. Of course, we do not have an access to natural numbers but the number k+1 may be represented by elements of $P_{k+1}-P_k$.

Here is the formal definition. Let $\rho(R,S,x,u,v)$ be the formula saying:
R(x,u,v), or $x\epsilon P_1$ & $(u,v)\epsilon Q_1$, or there is y such that R(y,y,y), $\Psi(R(y,\_,\_),S(y,\_,\_),u,v)$, S(y,x,x), and $\Psi(R(y,\_,\_),S(y,\_,\_),x,x)$.

Let $\sigma(R,S,x,u,v)$ be the formula saying:
S(x,u,v), or $x\epsilon P_1$ & $-[(u,v)\epsilon Q_1]$, or there is y such that R(y,y,y), $-\Psi(-S(y,\_,\_),-R(y,\_,\_),u,v)$, S(y,x,x), and $\Psi(R(y,\_,\_),S(y,\_,\_),x,x)$.

Here the expressions $x\epsilon P_1$ and $(u,v)\epsilon Q_1$ abbreviate pseudo first-order formulas $\varphi(\emptyset,x)$ and $\Psi(\emptyset,-\emptyset,u,v)$ respectively. Obviously, ρ and σ are positive in R and S. Therefore the operator H(R,S)=

$$(\{(x,u,v): \rho(R,S,x,u,v)\}, \{(x,u,v): \sigma(R,S,x,u,v)\}).$$

is monotone and has a least fixed point.

Lemma 8. The least fixed point of H is

$$(\cup_{k<m}[(P_{k+1}-P_k)\times Q_{k+1}],$$
$$\cup_{k<m}[(P_{k+1}-P_k)\times -Q_{k+1}]).$$

Proof. For each natural number k, let $(R_k,S_k)=H^k(\emptyset,\emptyset)$. It suffices to prove that

$$R_{k+1}-R_k=(P_{k+1}-P_k)\times Q_{k+1},$$
$$S_{k+1}-S_k=(P_{k+1}-P_k)\times -Q_{k+1}.$$

The case k=0 is clear: the formulas $\rho(\emptyset,\emptyset,x,u,v)$ and $\sigma(\emptyset,\emptyset,x,u,v)$ describe $P_1\times Q_1$ and $P_1\times -Q_1$ explicitly. Let k >0. If $(x,u,v)\epsilon (P_{k+1}-P_k)\times Q_{k+1}$ then any $y\epsilon P_k-P_{k-1}$ will witness that $\rho(R_k,S_k,x,u,v)$ & $-R_k(x,u,v)$ holds i.e. $(x,u,v)\epsilon R_{k+1}-R_k$. If $\rho(R_k,S_k,x,u,v)$ & $-R_k(x,u,v)$ holds, let y be any witness for ρ. By the inductive hypothesis, $y\epsilon R_i-R_{i-1}$ for some positive $i\leq k$, $R_k(y,\_,\_)=Q_i$, and $S_k(y,\_,\_)=-Q_i$. Hence $\Psi(Q_i,-Q_i,u,v) \leftrightarrow (u,v)\epsilon Q_{i+1}$, $S_k(y,x,x) \rightarrow -[x\epsilon P_i]$, and $\Psi(Q_i,-Q_i,x,x) \leftrightarrow (x,x)\epsilon Q_{i+1} \leftrightarrow x\epsilon P_{i+1}$; thus (x,u,v) belongs to $(P_{i+1}-P_i)\times Q_{i+1}$. But it does not belong to R_{k-1}. Hence i=k and $(x,u,v)\epsilon(P_{k+1}-P_k)\times Q_{k+1}$. The other equality is proved similarly. □

Theorem 1 follows from Lemma 8. □

Remark. To see the use of finiteness in the proof, note "any $y\epsilon P_k-P_{k-1}$ will witness" in the proof of Lemma 8.

Theorem 1 and Simultaneous Induction Lemma for FO+LFP imply Main Theirem.

§4. The collapse of the FO+IFP hierarchy

Again, all structures are supposed to be finite.

Theorem 1. Every FO+IFP formula is equivalent to an FO+IFP formula φ such that φ is either first-order or of the form $[\text{IFP}_{...}\Phi](...)$ where Φ is first-order.

Proof. Formulas φ, described in Theorem 1, will be called explicitly low. Also predicates $\{x: \Phi(x)\}$ and $IFP_{...}\Phi$, where Φ is first-order, will be called explicitly low. A formula (resp. predicate) will be called low if it is equivalent (resp. equal on every relevant finite structure) to an explicitly low formula (resp. predicate). We prove that every FO+IFP formula is low.

Lemma 1. Predicates $IFP1_{P,Q:x,y}(\varphi,\psi)$ and $IFP1_{P,Q,R:x,y,z}(\varphi,\psi,\chi)$, where φ, ψ and χ are first-order, are low.

Proof. Use simultaneous induction lemmas for FO+IFP. □

Lemma 2. If $\alpha(Q^*,x)$ is built from an explicitly low predicate $Q^*=IFP_{Q:y}\psi(Q,y)$ by first-order means then it is low. Hence, the set of low formulas is closed under negation and universal quantification.

Proof. $\alpha(\varphi,x)$ is equivalent to $[IFP1_{P,Q:x,y}(\Phi,\Psi)](x)$ where

$$\Phi(P,Q,x) = \forall x(\psi(Q,y)\leftrightarrow Q(y))\ \&\ \alpha(Q,x),$$
$$\Psi(P,Q,y) = \psi(Q,y).$$

The idea is: first build Q^*, then set $P=\{x: \alpha(Q^*,x)\}$. For readability, we have omitted a formally required disjunct $P(x)$ in the first clause. □

Lemma 3. The conjunction $[IFP_{P:x}\varphi](x)\ \&\ [IFP_{Q:y}\psi](y)$ of explicitly low formulas is equivalent to an explicitly low formula $[IFP1_{P,Q:x,y}(\Phi,\Psi)](x,y)$.

Proof.
$$\Phi(P,Q,x) = \varphi(P,x)\ \&\ \forall y[\psi(Q,y)\leftrightarrow Q(y)]\ \&\ Q(y),$$
$$\Psi(P,Q,y) = \psi(Q,y).\ □$$

Note. In Lemmas 2 and 3, x and y do not have common variables.

Lemma 4. If $\varphi(P,x)$ is an explicitly low formula $[IFP_{Q:y}\psi(P,Q,y)](t)$ then $IFP_{P:x}[P(x) \text{ or } \varphi(P,x)]$ is low.

Proof. Since y is bound in $IFP_{Q:y}\psi(P,Q,y)$, we can assume that no y-variable occurs in t or x. Note that if a t-variable does not occur in x then its value is not changed during the double recursion; view such variables as parameters.

Let $F(P)=\{x: \varphi(P,x)\}$, $P_i=F^i(\emptyset)$, and $G_i(Q)=\{y: \psi(P_i,Q,y)\}$. Let $m=\min\{i: P_i=P_{i+1}\}$ and, for every $i\leq m$, let $n_i=\min\{j: G_i{}^j(\emptyset)=G_i{}^{j+1}(\emptyset)\}$. We reduce the double induction to a single simultaneous induction. The difficulty is that Q oscillates, during the double recursion, between the empty set and the inductive fixed points of G_i's. To cope with this non-inflationary behavior, we introduce predicates $R(x,y)$ and $Q_0(y)$ where Q_0 is simply Q in its first incarnation. The stages of the simultaneous induction are labeled by pairs (i,j), where $i\leq m$ and $j\leq n_i$, ordered lexicographically. The dynamics of $P(x)$, $Q_0(y)$, and $R(x,y)$ is explained in Claims 1-4 below. Formally speaking, let

$$\alpha(x) = P(x) \text{ or } \forall y'(\psi(\emptyset,Q_0,y')\leftrightarrow Q_0(y'))\ \&\ Q_0(t) \text{ or } \forall y(\psi(P,Q,y)\leftrightarrow Q(y))\ \&\ Q(t);$$
$$\beta(y') = \psi(\emptyset,Q_0,y');$$

$$\gamma(x,y)= R(x,y), \text{ or } P(x)\ \&\ (\psi(P,Q,y), \text{ or } \forall y'(\psi(P,Q,y')\leftrightarrow Q(y'))\ \&\ \exists x(Q(t)\ \&\ -P(x));$$
where $Q(y)\leftrightarrow(\exists x'(Px'\ \&\ -\forall y'Rx'y'\ \&\ Rx'y) \text{ or } \exists x'P(x')\ \&\ \forall x'(Px'\rightarrow\forall y'Rx'y'))$.

Claim 1. On every stage (0,j), $P=R=\emptyset$ and $Q_0=G_0{}^j(\emptyset)$.

Claim 2. Suppose that $i>0$, $P=P_i$ and $R=(P_{i-1}\times All)\cup[(P_i-P_{i-1})\times G_i{}^j(\emptyset)]$. Then $Q=G_i{}^j(\emptyset)$. (Here "All" is the set of all tuples of the appropriate length.)

Claim 3. On every stage (i,j) with $i>0$, $P=P_i$; $Q_0=IFP(G_0)$; and $R=(P_{i-1}\times All)\cup[(P_i-P_{i-1})\times G_i{}^j(\emptyset)]$.

Claim 4. The simultaneous recursion stops on stage (m,n_m).

Thus $IFP_{P;x}\varphi(P,x)$ equals $IFP1_{P,Q,R;x,y',y}(\alpha,\beta,\gamma)$. □

The four lemmas imply Theorem 1. □

References:

AG M. Ajtai and Y. Gurevich, "Monotone versus Positive". To appear.

AU A. V. Aho and J. D. Ullman, "Universality of Data Retrieval Languages". 6th ACM Symp., Princip. of Prog. Languages, 1979, 110-120.

BGK A. Blass, Y. Gurevich and D. Kozen, "A Zero-One Law for a Logic with Fixed-Point Operator". Tech. Rep. CRL-TR-38-84, Sep 84, Comp. Research Lab, U. of Michigan.

CH A. Chandra and D. Harel, "Structure and Complexity of Relational Queries". 21st IEEE Symp., Found. of Comp. Sci, 1980, 333-347.

Gu Y. Gurevich, "Toward Logic Tailored for Computational Complexity". In "Computation and Proof Theory" (Editors M. M. Richter et al.), Springer Lecture Notes in Math. 1104 (1984), 175-216.

GS Y. Gurevich and S. Shelah, "Fixed-Point Extensions of First-Order Logic". Annals of Pure and Applied Logic, to appear.

Im1 N. Immerman, "Relational Queries Computable in Polynomial Time". 14th ACM Symp. on Theory of Computing, 1982, 147-152.

Im2 N. Immerman, "Relational Queries Computable in Polynomial Time". Manuscript, Yale University, January 1985.

Li A. Livchak, "The Relational Model for Process Control". Automatic Documentation and Math. Linguistics 4 (1983), 27-29.

Mo Y. N. Moschovakis, "Elementary Induction on Abstract Structures". North-Holland, 1974.

Ul J. D. Ullman, "Principles of Database Systems". Computer Science Press, 1982.

Va M. Vardi, "Relational Queries Computable in Polynomial Time". 14th ACM Symposium on Theory of Computing, 1982, 137-146.