Talk about generalised quantifiers

Thomas Forster

May 11, 2007

Contents

This is a file intended in the first instance for my own perusal. That is to say, it explains in great detail precisely those things that i think i might forget. It's not designed to be accessible to eavesdroppers, tho' i hope it is!

Chapter 1

Logic with the cofinite quantifier

CAPTAIN	... And I'm never, never sick at sea!
CHORUS	What, never?
CAPTAIN	No, never!
CHORUS	What, *never?*
CAPTAIN	*Hardly* ever!

Gilbert and Sullivan: HMS Pinafore

The logic that concerns me is the set of formulæ of LPC that are valid if we read $\forall$ as $\forall_\infty$ and $\exists$ as $\exists_\infty$ meaning 'for cofinitely many ... ' and 'for infinitely many ... ' respectively. $\forall_\infty x \neg$ is the 'hardly ever' quantifier.

I got interested in this logic for two reasons, both discussed below. (i) I noticed that a proof I gave years ago concerning $\exists^*$ and $\forall^*\exists^*$ formulæ and the term model for *NF*O could be cunningly refined to prove the analogous result for the same formulæ read in the infinitary way. (ii) It seemed to me that many facts about WQOs could be expressed neatly in the new language.

1.1 Stuff to fit in

- Slaman-Marker's [**?**] result on arithmetic in the cofinite Logic.

- Like quantifiers do not commute, but it might be that if $\forall_\infty y \forall_\infty x \Phi$ is valid so is $\forall_\infty x \forall_\infty y \Phi$. I'm pretty sure they commute if the variables are independent

- We should think very carefully about Skolemisation. Skolemisation is important beco's—*inter alia*—it will enable us to make sense of $\mathcal{L}^\infty$ with branching quantifiers. It seems we can Skolemise $\mathcal{L}^\infty$ easily enough. If 'x'

is bound by '$\exists_\infty$' and preceded by '$\forall_\infty y$' then replace all its occurrences by '$f(n, x)$' where 'n' is a variable ranging over naturals (or any other fixed countable set, and replace '$\exists_\infty x$' by '$\forall_\infty n \in \mathbb{N}$'; analogously with more universally bound vbls in earlier parts of the header. However you then have to add axioms for f to say things like $(\forall_\infty n \forall_\infty m)(f(x, m) \neq f(x, n))$. Two points arise here. (i) Torkel says his completeness theorem doesn't work for the language with function letters. (ii) we now seem to need equality after all, since the argument that equality can be removed depends on the equations being between variables not terms

- It's clearly not compact, beco's the set $\{\phi(a_i) : i \in \mathbb{N}\} \cup \{a_i \neq a_j : i \neq j\} \cup \{(\forall_\infty x)(\forall_\infty y)(\phi(x) \to \phi(y) \to x = y)\}$ is inconsistent but no contradiction can be deduced from it in finitely many steps.

- Since every formula is equivalent to one without equations, it ought to be possible to axiomatise the version without equality!

- Is it obvious that it is undecidable? This decidability business is potentially very interesting. Suppose this logic is decidable. Then consider the following. Take an interesting undecidable but axiomatisable theory T expressed in ordinary predicate calculus. Reinterpret the quantifiers so you now have a theory T^∞ in the new logic. T^∞ stands a very good chance of being decidable. What is certainly true is that the set of consequences in the new logic of the translations of an axiomatisation of T in LPC is decidable, but that isn't quite the same thing, but it still seems worth a look. I *would* like to know if its decidable!

- If ϕ is a sentence of this logic, then the class of models of ϕ is closed under "finite mutilation" (addition or removal of finitely many tuples from the extensions of predicates);

- Remember the very cute $\forall_\infty \exists_\infty$ version of $\leq^+$ and BQOs.

- If ϕ is a formula of LPC, let ϕ_n be the formula obtained from ϕ by reading $\forall$ as 'for all but at most n' and $\exists$ as 'there are at least n'. The set of valid formulæ of LPC is semidecidable, and for each n the set of formulæ of this form is decidable so, for each n, the set of ϕ such that ϕ_n is valid is axiomatisable.

- Is the correct notion of validity "true in all infinite models"? (Free-associate to Trakhtenbrot's theorem) Every existential sentence is false in all finite models ...

- Is it simpler if we drop equality from the language? One might expect it, but the sequent rules seem to make heavy use of equality!

- "For infinitely many $n \in \mathbb{N}$ there are infinitely many $k \in \mathbb{N}$ such that there are as many primes in $[k, k + n]$ as there are in $[0, n]$" expressed easily with the cofinite quantifier.

- Sometimes i wonder if it might be easier for first-years if one were to explain convergence of sequences by means of the cofinite quantifier: $\{a_n\}$ converges iff for all $x > 0$, for cofinitely many n and for cofinitely many m, $|a_n - a_m| < x$. One particularly nice feature is that use of the cofinite quantifier allows one to express things formally using the same number of quantifiers as English does.

- I think
$$(\forall_\infty x)(\forall_\infty y)(G(x) \wedge G(y) \to x = y)$$
sez there are only finitely many things which are G.

  What can go wrong? there might be finitely many x such that for each of them there are infinitely many y, all of them G, and all distinct from x. But this can't happen. Let x_0 be one of these finitely many bad xs, and let Y be an infinite set of witnesses—things that are G but nevertheless distinct from x_0. But then, for cofinitely many of these ys, there are only finitely many G-things not equal to y, so there really only are finitely many G-things after all.

  This is subtly different from the usual setting, and there may be some undiagnosed vicious misunderstanding in my mind. In the LPC case this formula says that the number of G things is small but not that it is null. In the new setting it says that it is null. Once we have a proof system for this logic it would be nice to see a proof!

- The lemma of Yasuhara about eliminabilty of equality raises an interesting point. The power of branching quantifiers seems to rely on exploitation of '='. What does this tell us about logic with branching cofinite quantifiers?

- It's pretty obvious that one cannot define ordinary quantifiers in terms of these new quantifiers. Should be easy to show that any formula with a finite model has arbitrarily large finite models, so there can be no formula true in a model with precisely one thing that is ϕ, so we cannot define the uniqueness quantifier, so we cannot define $\forall$ or $\exists$ either.

- Consider—for the moment—structures for this logic which have a binary $\in$ only. Directed edges. Now if we have a model $\mathcal{M}$ it should be elementarily equivalent (from the point of this language) to the result of deleting from it any edges according to the following scheme: if x is a vertex of finite indegree (outdegree), delete all edges going into (out of) it. Easy to see why one might expect this to work. But what about the following? Let x be a set that has infinitely many members, each of which belong to only finitely many things. We delete all the edges leaving them, but x now becomes empty, which wasn't part of the plan.

  Now, in order for this to change the truth value of a closed sentence it has to be something like $(\exists_\infty x)(\exists_\infty y)(y \in x \wedge (\forall_\infty z)(y \not\in x))$. This deletion can turn a model of this sentence into a structure that isn't a model of the sentence

- I think perhaps the best way to cope with non-commutation of like quantifiers to to deem all blocks of like quantifiers to be branching. This might bugger up the PNF theorem. Must get straight how it can be that (as Yasuhara claims) the PNF thm holds for this logic but like quantifiers do not commute.

  The answer i think is that if we coerce $(\forall_\infty x)(F(x)) \wedge (\forall_\infty y)(G(y))$ into PNF we get either
  $$(\forall_\infty x)(\forall_\infty y)(F(x) \wedge G(y))$$
  or
  $$(\forall_\infty y)(\forall_\infty x)(F(x) \wedge G(y))$$
  but it deosn't matter becos 'x' and 'y' are not connected so the quantifiers commute.

1.2 Yasuhara

Yasuhara considers something slightly more general, namely $\exists_\infty$ where ∞ is some arbitrary infinite cardinal, not necessarily $\aleph_0$. ("There are at least ∞ things such that ...") He notes that the following hold

$$\exists_\infty x A \longleftrightarrow \exists_\infty y A[y/x]$$

(renaming of bound variables

$$\exists_\infty x(A \wedge B) \longleftrightarrow (\exists_\infty x A) \wedge B$$

('x' not free in B)

$$\exists_\infty x(A \vee B) \longleftrightarrow (\exists_\infty x A) \vee B$$

('x' not free in B)

$$\exists_\infty x(A \vee B) \longleftrightarrow (\exists_\infty x A \vee \exists_\infty x B)$$

$$\forall_\infty x(A \wedge B) \longleftrightarrow (\exists_\infty x A \wedge \exists_\infty x B)$$

and he notes the following equality axioms
$\neg(\exists_\infty x)(a = x \wedge A(x))$
$\neg(\exists_\infty x)(x \neq x \wedge A(x))$
$(\forall_\infty x)(a = x \vee A(x))$
$(\forall_\infty x)(x = x \vee A(x))$
$(\exists_\infty x)(a \neq x \wedge A(x)) \longleftrightarrow (\exists_\infty x)(A(x))$
$(\exists_\infty x)(x = x \wedge A(x)) \longleftrightarrow (\exists_\infty x)(A(x))$
$(\forall_\infty x)(a = x \vee A(x)) \longleftrightarrow (\exists_\infty x)(A(x))$
$(\forall_\infty x)(x \neq x \vee A(x)) \longleftrightarrow (\exists_\infty x)(A(x))$

He points out that there is a Prenex normal form theorem, in the sense that any formula A can be transformed into something A' in PNF s.t. the biconditional $A \longleftrightarrow A'$ is (semantically) valid.

Let's have a proof of this at some point.

He comments without proof that

LEMMA 1. *Every formula is equivalent to one wherein there is no '$=$' within the scope of a quantifier.*

Proof: (Exercise for tf): We perform an induction on the subformula relation.

For quantifier-free formulæ there is nothing to prove. Induction step for connectives is obvious. Let us consider the induction step for $\exists_\infty$.

We have a formula $(\exists_\infty x)\Phi$ where Φ is a boolean combination of formulæ wherein there is no '$=$' within the scope of a quantifier. There may be occurrences of '$=$' within Φ of course. If neither of the two variables in the equation are x then the equation can be pulled outside the scope of the '$\exists_\infty$', thereby satisfying the induction hypothesis. There remains the case where one of the variables is x. WLOG we can assume that Φ is in disjunctive normal form, in the sense that it is a disjunction of conjunctions of quantified formulæ containing no equations, and of atomics and negatomics. We are concerned about disjuncts that contain equations '$x = a$' or their negations. Any disjunct $x = a \wedge \psi$ can be replaced by $\psi[a/x]$ thereby removing the equation. All these manœuvres are standard in LPC: it is the next move that is peculiar to the logic of the cofinite quantifier.

In the case of an inequation we are considering $(\exists_\infty x)((x \neq a \wedge \psi) \vee \phi)$

This is equivalent to $(\exists_\infty x)(x \neq a \wedge \psi) \vee (\exists_\infty x)\phi)$

and $(\exists_\infty x)(x \neq a \wedge \psi)$ is equivalent to

$(\exists_\infty x)(\psi)$

The proof for the universal quantifier is presumably dual; I haven't tho'rt about it. ■

(Digression for NFistes only. This process of eliminating equality can turn an unstratified formula into a stratified one: consider $(\exists_\infty y)(y \in x \wedge y \neq x)$.)

1.3 A sequent formulation of the logic of the cofinite quantifier

I use semicolons for separating sequents, and commas for separating Gammas and Deltas as usual.

One obvious approach is to introduce rules for $\forall_n$ and $\exists_n$ for every n and build the rules for $\forall_\infty$ and $\exists_\infty$ from them. Thus, for each n, we have a rule

$$\frac{\Gamma \vdash \Delta, (\forall_n x)F(x)}{\Gamma \vdash \Delta, (\forall_\infty x)F(x)}$$

and, for each n, we have a rule

$$\frac{\Gamma, (\exists_n x)F(x) \vdash \Delta}{\Gamma, (\exists_\infty x)F(x) \vdash \Delta}$$

and two infinitary rules. Infer

$\Gamma \vdash \Delta, (\exists_\infty x)F(x)$ once you have, for each n, $\Gamma \vdash \Delta, (\exists_n x)F(x)$

and infer

$\Gamma, (\forall_\infty x)F(x) \vdash \Delta$ once you have, for each n, $\Gamma, (\forall_n x)F(x) \vdash \Delta$

Or we could go at it bald-headed:

$$\frac{\Gamma \vdash \Delta,\ \phi(x),\ x = t_1,\ \ldots\ x = t_n \ldots t_i = t_j \ldots}{\Gamma \vdash \Delta,\ (\forall_\infty x)(\phi(x))} \qquad (\forall_\infty \text{ -R})$$

$$\frac{\Gamma,\ \phi(x) \vdash \Delta,\ x = t_1,\ \ldots\ x = t_n \ldots t_i = t_j \ldots}{\Gamma,\ (\exists_\infty x)(\phi(x)) \vdash \Delta} \qquad (\exists_\infty\text{-L})$$

In both rules 'x' must not be free in the lower sequents; in the upper sequents n can be anything in $\mathbb{N}$, and the 'i' and 'j' range over naturals up to n.

The other pair of rules are messier.

Clearly the rule for $\exists_\infty$-R will have a precedent consisting of $\Gamma \vdash \Delta, \phi(t_i)$ for each i in some infinite set I. The succedent is the single sequent $\Gamma \vdash \Delta, (\exists_\infty x)(\phi(x))$

$$\frac{\Gamma \vdash \Delta, \phi(t_0); \ldots \Gamma \vdash \Delta, \phi(t_n); \ldots \Gamma \vdash \Delta, t_i \neq t_j \ldots}{\Gamma \vdash \Delta, (\exists_n x)(\phi(x))} \quad \exists_\infty\text{-R}$$

and

The rule $\forall_\infty$-L will dually have (not a precedent with infinitely many sequents but) a precedent with an infinite sequent.

$$\frac{\Gamma, \phi(t_0) \ldots \phi(t_n) \vdash \Delta, \ldots t_i = t_j \ldots}{\Gamma, (\forall_n x)(\phi(x)) \vdash \Delta} \forall_\infty\text{-L}$$

for all $i \neq j$ both $< n$.

Let's have a brief reality check. Must make sure these quantifiers are dual in the usual way:

$$\cfrac{\cfrac{\vdash F(a), \neg F(a)}{\vdash F(a), (\exists x)(\neg F(x))}}{\vdash (\forall x)(F(x)), (\exists x)(\neg F(x))}$$

So in trying to prove $\vdash (\forall_n x)(F(x)), (\exists_n x)(\neg F(x))$
we should expect to attack the $\forall_n$ first, thus:

$$\frac{\vdash (\exists_n x)(\neg F(x)), F(x),\ x = t_1,\ \ldots\ x = t_n \ldots t_i = t_j \ldots}{\vdash (\exists_n x)(\neg F(x)),\ (\forall_n x)(F(x))}$$

Now we attack $(\exists_n x)(\neg F(x))$. Abbreviate '$F(x),\ x = t_1,\ \ldots\ x = t_n \ldots t_i = t_j \ldots$' to '$\Delta$' getting

$$\frac{\vdash \Delta, \neg F(t_0), \ldots \vdash \Delta, \neg F(t_n), \ldots \vdash \Delta, t_i \neq t_i \ldots}{\vdash \Delta, (\exists_n x)(\neg F(x))}$$

Let's have a brief reality check. Let's try to prove $\neg(\exists_\infty x)\neg\phi(x) \vdash (\forall_\infty x)(\phi(x))$. This obviously came from

$\vdash (\forall_\infty x)(\phi(x)), (\exists_\infty x)\neg\phi(x)$.

This must have come from

$$\frac{\ldots \vdash (\forall_\infty x)(\phi(x)), (\exists_n x)\neg\phi(x) \ldots}{\vdash (\forall_\infty x)(\phi(x)), (\exists_\infty x)\neg\phi(x)}$$

Torkel's challenge: prove

$$(\exists_\infty x)(P(x) \wedge Q(x)) \vdash (\exists_\infty x)P(x)$$

My challenge: prove

$$(\forall_\infty x)(\forall_\infty y)(G(x) \wedge G(y) \to x = y) \to \neg(\exists_\infty x)(G(x))$$

1.4 Torkel's Completeness Proof

Your first guilty tho'rt might be that the alleged completeness of this logic might contradict Lindström's theorem, if only you could remember what it was! Lindström's theorem says that there is no logic stronger than ordinary LPC which is compact and has downward Skolemheim. Doesn't say anything about completeness/axiomatisability

1.4.1 Axiomatisation

The system has as its only quantifier $\exists_\infty x$, meaning "there are infinitely many x such that". We have as rules some suitable formulation of the usual rules for propositional logic and identity, and the specific rules for $\exists_\infty$:

1) All instances of ordinary propositional axioms and identity axioms,

2) the further axioms (for every A and B)

1. $(\exists_\infty x)A \to A$ if x is not free in A;
2. $(\exists_\infty x)A(x) \to (\exists_\infty y)A(y)$;
3. $\neg(\exists_\infty x)(x = s)$;
4. $(\exists_\infty x)(A \vee B) \to (\exists_\infty x)A \vee (\exists_\infty x)B$;
5. $(\exists_\infty x)A(x) \to (\exists_\infty x)(\exists_\infty y)(A(x) \wedge A(y) \wedge \neg x = y)$.

We have the rule: if $A \to B$ has been proved, conclude $(\exists_\infty x)A \to (\exists_\infty x)B$".

We need three little lemmas

Lemma 2. *(The constant elimination property) If $G(c) \vdash A(c)$ where c is a constant, $G(x) \vdash A(x)$*

Lemma 3. *For every m, every G, every A and B: If $G \vdash A(x_1) \wedge \ldots \wedge A(x_m)$ and all the x_i are different $\rightarrow B(x_1) \vee \ldots \vee B(x_m)$ and $G \vdash \exists_\infty x A(x)$ then $G \vdash \exists_\infty x B(x)$.*

Lemma 4. *For every n, every $A_1, \ldots A_n$ and every B:*
If

$$G \vdash A_1(x_1) \wedge \ldots \wedge A_n(x_n) \rightarrow B(x_1) \vee \ldots \vee B(x_n)$$

then

$$G \vdash \exists_\infty x A_1(x) \wedge \ldots \wedge \exists_\infty x A_n(x) \rightarrow \exists_\infty x B(x)$$

This seems to go nicely as a sequent rule:

$$\frac{\Gamma, A_1(x_1), \ldots A_n(x_n) \ \vdash \ \Delta, B(x_1), \ldots B(x_n)}{\Gamma, \exists_\infty x A_1(x), \ldots \exists_\infty x A_n(x) \ \vdash \ \Delta, \exists_\infty x B(x)}$$

We claim that these three lemmas follow from the rules given. Note that even if there is some oversight in my verification of this, this is not an essential difficulty, since they are clearly semantically sound, and so we would just add whatever is needed to get them

1.4.2 A Henkin-style Completeness proof

Theorem 5. *Every consistent formula has a model*

Proof:
Let G_0 be $\{A\}$, where A is consistent. We assume that A contains no function symbols, though it may contain individual constants.

We have an infinite supply of special constants, individual constants that do not occur in A. We have an enumeration $A_0, A_1, A_2, \ldots$ of all formulas in the language obtained by adding all the special constants.

We define G_{n+1} by recursion as usual.

1. If A_n is not of the form $\exists_\infty x B(x)$ or $\neg\exists_\infty x B(x)$:

   $G_{n+1} = G_n \cup \{A_n\}$ if A_n is consistent with G_n, otherwise $G_{n+1} = G_n \cup \{\neg A_n\}$.

2. If A_n is of the form $\exists_\infty x B(x)$ or $\neg\exists_\infty x B(x)$:

   (a) $\exists_\infty x B(x)$ is consistent with G_n. In this case we add $\exists_\infty x B(x)$, and we also add an infinite number of sentences $B(c_1), B(c_2), \ldots$ where the c_i are individual constants that have not been previously used for this purpose, and do not occur in any of the formulas $A_0, \ldots, A_n$, and we also add $c_i \neq c_j$ for different i, j, forcing these constants to be interpreted as different individuals. We leave infinitely many special constants still unused.

(b) $\exists_\infty x B(x)$ is inconsistent with G_n.

i. $\neg B(x)$ is inconsistent with G_n. (Example: G_0 may consistently contain $\neg\exists_\infty x(x = x)$.) In this case we only add $\neg\exists_\infty x B(x)$ to obtain G_{n+1}.

ii. $\neg B(x)$ is consistent with G_n. In this case we add $\neg\exists_\infty x B(x)$, but also (since we want to force the interpretation to be such that only finitely many x satisfy $B(x)$) we add $\neg B(c)$ for *every* special constant c which does not occur in any of the formulas $A_0, \ldots, A_n$.

G is of course the union of the G_n.

Now suppose we manage to prove G consistent. Then we define, as usual, an interpretation whose individuals are the equivalence classes of constants, with constants a and b equivalent if $a = b$ belongs to G. This yields a model, by the usual reasoning. In the crucial case, where A is $\exists_\infty x B(x)$, we need to verify that A is in G if and only if $B(a)$ is in G for infinitely many constants a, and indeed for infinitely many equivalence classes $[a]$. In one direction, this is immediate: if $\exists_\infty x B(x)$ is in A, we've added all the $B(c)$ and inequalities between the c. In the other direction, if $\exists_\infty x B(x)$ is not in G, we've added at some stage $\neg\exists_\infty x B(x)$, and then also added $\neg B(c)$ for all but finitely many constants, so there can't be infinitely many c, let alone infinitely many equivalence classes, such that $B(c)$ is in G.

So the whole trick is now to prove G consistent. Suppose G_{n+1} is inconsistent. We have two cases.

1. case 1: We added $\exists_\infty x B(x)$, and we also added $B(c_1)$, $B(c_2)$, $\ldots$ and inequalities between the newly added constants. Suppose a proof of an inconsistency uses $c_1, \ldots c_n$. Now the difference between the present setting and the usual Henkin proof for standard logic is that the newly added $c_1, c_2 \ldots$, even though they have not been used before for this purpose, is that the $c_1, c_2, \ldots$ may well occur in formulas that have been added before, namely in earlier case-2 situations. So there may be formulas

$$\begin{array}{ccc} \neg A_1(c_1), & \ldots, & \neg A_m(c_1) \\ \neg A_1(c_2), & \ldots, & \neg A_m(c_2) \\ \vdots & \ldots & \vdots \\ \neg A_1(c_n), & \ldots, & \neg A_m(c_n) \end{array}$$

among those earlier added. But these are the only possible occurrences of the c_i in earlier formulas. So there is a derivation of an inconsistency from

$\neg A_1(c_1), \ldots, \neg A_m(c_1), \neg A_1(c_2), \ldots, \neg A_m(c_2), \ldots..\neg A_1(c_n), .., \neg A_m(c_n)$, all the c_i different, $B(c_1), B(c_2), \ldots B(c_n)$.

But then, by the constant elimination property, there is a derivation of

$B(x_1) \wedge \ldots \wedge B(x_n)$ and all the x_i different $\rightarrow C(x_1) \vee \ldots \vee C(x_n)$

where $C(x)$ is $A_1(x) \vee \ldots \vee A_m(x)$. So by property 1 above, and the fact that $\exists_\infty x B(x)$ is derivable, it follows that $\exists_\infty x C(x)$ is derivable, and therefore the disjunction of $\exists_\infty x A_i(x)$, contradicting the consistency of earlier sets in the sequence and the provability of $\neg\exists_\infty x A_i(x)$ for every x.

2. We added $\neg\exists_\infty x B(x)$, and also added $\neg B(c)$ for every special constant not occurring in any of $A_0 \ldots, A_n$. An inconsistency means then that G_n proves the disjunction $B(c_1) \vee \ldots \vee B(c_n)$ for some special constants $c_1 \ldots, c_n$.

   First, if there are no previous Qx-additions to the set, no c_i occurs in G_n, so G_n proves $B(x_1) \vee \ldots \vee B(x_n)$, so G_n proves $B(x)$, and we have in fact the subcase in which only $\neg\exists_\infty x B(x)$ is added. So there must be some previous $\exists_\infty x$-additions to the set. The reasoning is now similar to that in case 1. We find that a subset of G_n must prove

$$A_1(x_1) \wedge \ldots \wedge A_n(x_n) \rightarrow B(x_1) \vee \ldots \vee B(x_n)$$

   where G_n also proves $\exists_\infty x A_1(x) \ldots, \exists_\infty x A_n(x)$, and the property 2 yields a contradiction.

If a sentence A in this logic is valid, it is provable using the above rules, but only if we do not allow function symbols in the language. If we do allow function symbols, we get such valid formulas as

$$(\exists_\infty x)p(f(x)) \wedge \neg(\exists_\infty x)p(x) \rightarrow (\exists_\infty x)(\exists_\infty y)(x \neq y \wedge f(x) = f(y))$$

which are not provable from the above axioms. We can prove this formula if we add the schema

$$(\exists_\infty x)(\exists_\infty y)(A(t(x)) \wedge A(t(y)) \wedge \neg(t(x) = t(y))) \rightarrow (\exists_\infty x)A(x)$$

(for any term $t(x)$), but I've failed completely to prove any completeness theorem for this or any other extension of the axioms in the case where we allow function symbols.

Does anybody have any ideas about this, or about any treatment of it in the literature? I haven't been able to find any such treatment.

Notice that altho' adjacent occurences of '$\forall$' commute, adjacent occurrences of '$\forall_\infty$' do not. I think that what this is trying to tell us is that when we come to consider formulæ in the cofinite logic that correspond to a formula of LPC with adjacent like quantifiers, such as:

$$(\forall x)(\forall y)(\forall z)(x \leq y \leq z \rightarrow x \leq z)$$

what we really want is something like

$$\begin{pmatrix} \forall_\infty x \\ \forall_\infty y \\ \forall_\infty z \end{pmatrix} (x \leq y \leq z \rightarrow x \leq z) \tag{1.1}$$

So let's think a bit about branching quantifiers in this connection

$$\begin{pmatrix} \forall_\infty y_1 \\ \forall_\infty y_2 \end{pmatrix} \phi(y_1, y_2)$$

This is really

$$\begin{pmatrix} (\exists Y_1 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_1 \in Y_1) \\ (\exists Y_2 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_2 \in Y_2) \end{pmatrix} \phi(y_1 y_2)$$

Skolemizing and flattening will give us

$$(\exists Y_1 Y_2)(\forall y_1 \in Y_1)(\forall y_2 \in Y_2)\phi(y_1 y_2)$$

where the Ys are variables over cofinite sets. This is emphatically *not* the same as either

$$(\forall_\infty y_1)(\forall_\infty y_2)\phi(y_1, y_2)$$

or

$$(\forall_\infty y_2)(\forall_\infty y_1)\phi(y_1, y_2)$$

...which are of course distinct from each other.

1.5 Nice embeddings

Let us say that an embedding f from a finite structure $\mathcal{M}_1$ into another struture $\mathcal{M}_2$ is **nice** if for every extension of $\mathcal{M}_1$ by the addition of one point there are infinitely many nice embeddings from the new structure into $\mathcal{M}_2$.

There is an obvious notion of a **nice family** $\mathcal{F}$ of embeddings from finite substructures[1] of $\mathcal{M}_1$ into $\mathcal{M}_2$: $\mathcal{F}$ contains the empty map and every injection i in $\mathcal{F}$ is a map from a finite substructure $\mathcal{M}'$ of $\mathcal{M}_1$ which, for any $m \in M_1 \setminus M'$, can be extended in infinitely many ways to another map in $\mathcal{F}$ whose domain is $M \cup \{m\}$.

We will need a special symbol to say that there is a nice family of maps from $\mathcal{M}_1$ to M_2. $\mathcal{M}_1 \prec_{\text{nice}} \mathcal{M}_2$, say.

For any countable graph there is a nice family of embeddings into the canonical random graph. I'm equally pretty sure that something similar happens with all countably categorical structures. It's certainly the case that for any countable total order there is a nice family of embeddings into the rationals.

Let G be the canonical countable random graph, and H another graph of which G is a subgraph ("induced subgraph"). It's obvious that (i) every existential sentence true in H is true in G. It's easy to show further that (ii) every universal-existential sentence true in H is also true in G. The analogues of (i) and (ii) also hold for sentences of the appropriate quantifier class in $\mathcal{L}(\infty)$.

[1]strictly, structures on finite subsets of M_1, the carrier set of $\mathcal{M}$.

1.6 Applications to the Quine systems

1.6.1 The term model of NF0

The term model of *NF*O can be thought of as the algebra of all words in *NF*O operations reduced by *NF*O-provable equations. This quotient is unique and well-defined; a proof can be found on p. 376 of [5]. It behaves in some ways like a countably categorical structure.

Theorem 6. *For every countable binary structure $\mathcal{M}$ there is a nice family of embeddings into the term model for NFO.*

Proof:

We will prove this by refining the construction of my 1987 paper to obtain a construction of a nice family of embeddings.

The 1987 construction takes a countable binary structure $\mathcal{M} = \langle M, R\rangle$ equipped with a wellordering of length ω and gives to each initial segment (or more strictly, its domain) an injection into the term model. We will do something slightly more complicated. We will not be providing injections-into-the-term-model to (domains of) initial segments of a fixed wellordering: our injections-into-the-term-model will be defined on the domains of finite partial functions from M to $\mathbb{N}$. We will think of these finite partial functions as lists of ordered pairs so that we can construct the nice family of injections by primitive recursion on lists. Doubled colons is our notation for consing things onto the front of lists, so that—to take a pertinant example—$\langle x, k\rangle :: s$ is the finite map that agrees with s on its domain and additionally sends x to k. We will construct for each s an injective homomorphism i_s from $dom(s) \in$ term-model-for-*NF*O, and this family of maps will be nice.

We will need an infinite supply of distinct selfmembered sets and an infinite supply of distinct non-selfmembered sets: such a supply can easily be found with the help of the B function. Let the nth *left* object be $B^n(V)$ and the nth *right* object be $B^n(\emptyset)$. All left objects are self-membered and no right objects are. The exponent gives us a convenient notion of *rank* of these left and right objects. It will be important in what follows that every value of any i_s has finite symmetric difference with a left object or a right object. It will also be important that any two left or right objects have infinite symmetric difference.

For s a finite partial map $M \to \mathbb{N}$ we will construct i_s from $dom(s)$ to the term model by primitive recursion on lists.

We start with the empty map from the empty substructure (the domain of the empty partial map).

The variable 's' will range over finite partial maps $M \to \mathbb{N}$ and for each s, i_s will be an injective homomorphism from $dom(s)$ to the term model for *NF*O.

For the recursion (primitive recursion on lists) let us suppose we have constructed a map i_s and we want to construct $i_{\langle x,k\rangle::s}$. And we must have $i_{s'} \neq i_{s''}$ whenever $s \neq s''$.

The construction of $i_{\langle x,k\rangle::s}$ from i_s is uniform in x and k. $i_{\langle x,k\rangle::s}$ will agree with i_s on $dom(s)$ of course. During the earlier construction of i_s we will have used some left objects and some right objects. Let n_s be the least n such that the only left or right objects touched so far in the construction of i_s have indices below n. Now, given $k \in \mathbb{N}$, we want X to be a left object or a right object, depending on whether $\mathcal{M} \models R(x,x)$ or not, and we set it to be the $(n_s + k)$th such object, or the $(n_s + k + 1)$th, if $(n_s + k)$ is odd. X is thus a left or right object, with a subscript that is even and is larger than any subscript we have seen so far.

$i_{\langle x,k\rangle::s}(x)$ will be obtained from X by adding and removing only finitely many things. We have to add things in A and delete things that are in B:

A: $\{i_s(m) : m \in dom(s) \ \wedge \mathcal{M} \models R(m,x)\}$

B: $\{i_s(m) : m \in dom(s) \ \wedge \mathcal{M} \models \neg R(m,x)\}$

C and D are harder to deal with:

C: $\{i_s(m) : m \in dom(s) \ \wedge \ \mathcal{M} \models R(x,m)\}$

E: $\{i_s(m) : m \in dom(s) \ \wedge \ \mathcal{M} \models \neg R(x,m)\}$

Our final choice for $i_{\langle x,k\rangle::s}(x)$ must extend A, be disjoint from B, belong to everything in C, and to nothing in E. There is no guarantee that X will do, but it's a point of departure; our first approximation to $i_{\langle x,k\rangle::s}(x)$ is $(X \setminus B) \cup A$.

For each m in $dom(s)$ let X_m be that left or right object from which $i_s(m)$ was obtained by the finite tweaking that we are about to explain. We want to control the truth-value of $i_{\langle x,k\rangle::s}(x) \in i_s(m)$. It's hard to see how to do this directly, but one thing we *can* control is the truth-value of $i_{\langle x,k\rangle::s}(x) \in X_m$, because this is the same as the truth-value of $B^{-1}X_m \in i_{\langle x,k\rangle::s}(m)$ and we can easily add or delete the various $B^{-1}(X_m)$ from $(X \setminus B) \cup A$.

Suppose for some particular m we want to arrange that $i_{\langle x,k\rangle::s}(x) \in i_s(m)$. We put $B^{-1}(X_m)$ into $i_{\langle x,k\rangle::s}(x)$. This ensures that $i_{\langle x,k\rangle::s}(x) \in X_m$. This is very nearly what we want, since the symmetric difference $X_m \ \Delta \ i_s(m)$ is finite. Now because we chose n_s to be larger than the subscript on any left or right object we had used so far in building i_s we can be sure that $i_{\langle x,k\rangle::s}(x)$ is not one of the finitely many things in $X_m \ \Delta \ i_s(m)$. So $i_{\langle x,k\rangle::s}(x) \in X_m$ and $i_{\langle x,k\rangle::s}(x) \in i_s(m)$ have the same truth-value.

In the light of this, we obtain $i_{\langle x,k\rangle::s}(x)$ from our first approximation—$(X \setminus B) \cup A$—by adding everything in $\{B^{-1}(X_m) : \mathcal{M} \models R(x,m)\}$ and deleting everything in $\{B^{-1}(X_m) : \mathcal{M} \models \neg R(x,m)\}$. Just a final check to ensure that this doesn't interfere with the adding and deleting we did initially, by adding everything in A and deleting everything in B: this last stage adds and deletes left-or-right objects with *odd* subscripts, whereas the initial tweaking added and deleted left-or-right objects (if any) with even subscripts only. ■

Corollary 7. *Every countable binary structure embeds into the term model of NFO in* $2^{\aleph_0}$ *ways.*

The general theme of this note is extending to the logic of the cofinite quantifier the various known results about ordinary logic and the Quine systems. We know that every $\exists^*$ sentence consistent with *NF*O holds in the term model. To get a version for the cofinite quantifier we need to get straight the idea of a $\exists^*_\infty$ formula consistent with *NF*O.

"Being consistent" in this sense for a formula $(\exists_\infty x_1 \ldots x_n)\phi$ where ϕ is quantifier free means the following. Suppose ϕ has n free variables. Then we invent constants whose suffixes come from $\mathbb{N}^{\leq n}$. For each sequence $c_{i_1} \ldots c_{i_n}$ where the suffix i_{k+1} is of length $k+1$ and is an end-extension of the suffix i_k, we adopt the axiom $\phi(c_{i_1} \ldots c_{i_n})$. Call this theory T. Then T is equivalent to $(\exists_\infty x_1 \ldots x_n)\phi$ in the sense that every model of T is an expansion of a model of $(\exists_\infty x_1 \ldots x_n)\phi$ and *vice versa*.

THEOREM 8. *Every $\exists^*_\infty$ formula consistent with NFO is true in all models of NFO.*

Proof: Let $(\exists_\infty x_1 \ldots x_n)\phi$ be such a formula, and T the theory obtained from it as above. Now every axiom of T is a consistent $\exists^*$ formula, and so is true in the term model, and so is a theorem of *NF*O.

■

Notice that we haven't yet had to exploit the clever construction of nice embeddings. That happens next.

REMARK 9. *The term model for NFO satisfies every $\forall^*_\infty \exists^*_\infty$ formula consistent with NFO.*

Proof: Consider $(\forall_\infty x_1 \ldots x_n)(\exists_\infty y_1 \ldots y_k)\phi(\vec{x}, \vec{y})$. Suppose this has a model $\mathcal{M}$. We want to show that it is true in the term model. For this it will suffice to show that if $\vec{t}$ is any tuple of terms such that $\mathcal{M} \models (\exists_\infty y_1 \ldots y_k)\phi(\vec{t}, \vec{y})$ then there are infinitely many many *terms* s_1 such that there are infinitely many terms s_2 etc such that $\phi(\vec{t}, \vec{s})$.

The first step is to simplify $(\exists_\infty y_1 \ldots y_k)\phi(\vec{t}, \vec{y})$ to the limits of our ingenuity. We know that atomic formulæ in ϕ need never be of the form '$y_j \in t_i$', because any such atomic wff can be expanded until it becomes a boolean combination of atomic wffs like '$y_i = t_j$', '$y_j \in y_i$', and '$t_j \in y_i$'. Then we can recast the matrix into disjunctive normal form. We know that $\mathcal{M} \models (\forall_\infty \vec{x})(\exists_\infty \vec{y})(\Phi(\vec{x}, \vec{y}))$ so there is at least one disjunct that does not trivially violate the theory of identity. This disjunct is a conjunction of things like '$y_i = t_j$', '$y_j \in y_i$', and '$t_j \in y_i$' and their negations, atomic wffs not containing any $\vec{y}$ having vanished since they are decidable.

We now have to find ways of substituting *NFO* terms $\vec{w}$ for the $\vec{y}$ to make every conjunct in the disjunct true. To do this we return to the constructions seen in the proof of theorem 1.6.1. We construct witnesses for the $\vec{y}$ in the way we constructed values of the function i in the proof of theorem 1.6.1. Let n_0 be some fixed integer such that all the t_i that appear in our disjunct have Bs nested less deeply than n. We know of (the infinitely many witnesses that we have to find for) y_0 that they is to have certain ts as members and certain others not. For each $k \in \mathbb{N}$ we construct a word w_0 which is the $n_0 + k$th left member (if '$y_0 \in y_0$' is a conjunct) or the n_0th right object (otherwise) $\cup$ (the tuple of

t_i such that '$t_i \in y_0$' is a conjunct) minus (the tuple of t_j such that '$t_j \notin y_0$' is a conjunct). From here on, we construct words w_i to be witnesses for y_i in exactly the same way as we proved theorem 1.6.1. ■

THEOREM 10. *If NFO* $\vdash \exists\vec{x}\forall\vec{y}\phi(\vec{x},\vec{y})$ *where* ϕ *is quantifier-free then for some tuple* $\vec{t}$ *of NFO words, we have NFO* $\vdash \forall\vec{y}\phi(\vec{t},\vec{y})$.

Proof: Let $\exists\vec{x}\forall\vec{y}\phi(\vec{x},\vec{y})$ be a $\exists^*\forall^*$ sentence, and suppose that for every tuple $\vec{t}$ of *NF*O terms it is consistent that the tuple $\vec{t}$ is not a witness to the $\vec{x}$. Then the scheme

$$(\exists\vec{y})(\neg\phi(\vec{t},\vec{y})) \text{ over all tuples of terms } \vec{t} \tag{1.1}$$

is consistent.

How complicated is scheme 1.1? Well, each instance is equivalent to a disjunction of things of the form $(\exists\vec{y})(\psi(\vec{t},\vec{y}))$ where ψ is a conjunction of atomics and negatomics. What sort of atomics and negatomics? Well, equations and inequations between the ts disappear beco's they are all T or F by elementary means. Equations $y = t$ can be removed by replacing all occurrences of 'y' by 't'. What's left? Inequations $y \neq t$ and $y \in t$, $t \in y$, $y \notin t$, $t \notin y$. We attack those recursively. $y \in t$ might be $y \in t_1 \wedge y \in t_2$, in which case we recurse further. If it is $y \in t_1 \vee y \in t_2$ then the $\exists^*$ formula in which it occurs gets split into two such formulæ. If we keep on doing this we will end up with a disjunction of $\exists^*$ formulae with terms appearing, but only in inequations or to the left of an '$\in$'. Clearly any such disjunction, if satisfiable at all, is satisfiable with the witnesses being finite tuples of terms, and is therefore true in the term model. So each instance of scheme 1.1 is true in the term model. That is to say, the term model believes $(\forall\vec{t})(\exists\vec{y})(\neg\phi(\vec{t},\vec{y}))$. So the original $\exists^*\forall^*$ sentence is not true in the term model, contradicting our assumption that *NFO* $\vdash \exists\vec{x}\forall\vec{y}\phi(\vec{x},\vec{y})$.

So if *NF*O proves a $\exists^*\forall^*$ sentence, there are provably witnesses that are *NF*O terms. ■

By now the reader will have thought enough about extending these results to isomorphic formulæ in the language with the cofinite quantifier to have spotted that in the last para of the last proof there are of course *infinitely many* ways of satisfying such disjunctions. Accordingly I hope that later draughts of this note will contain a proof of

THEOREM 11. *If NFO* $\vdash (\exists_\infty\vec{x})(\forall_\infty\vec{y})\phi(\vec{x},\vec{y})$ *where* ϕ *is quantifier-free then for a suitable infinity of tuples* $\vec{t}$ *of NFO words, we have NFO* $\vdash (\forall_\infty\vec{y})\phi(\vec{t},\vec{y})$.

We must think a bit about the scenario that the theorem describes. "*NFO* $\vdash (\exists_\infty\vec{x})(\forall_\infty\vec{y})\phi(\vec{x},\vec{y})$" means simply that in every model of *NF*O we can find infinitely many x_1 such that for each of them we can find infinitely many x_2 etc. The claim then is that, whenever this happens, we can take this network of xs to be *NF*O terms.

Now suppose the claim is false, and that altho' in every model of *NF*O we can find infinitely many x_1 such that for each of them we can find infinitely many x_2 etc., we cannot take all of these witnesses to be terms.

That is to say, if we take any set of countably many terms, and think of them as t_s where s is a sequence of natural numbers of length at most the length of $\vec{x}$, then the scheme

$$(\forall_\infty \vec{y})\phi(t_i, t_{i,j}, t_{i,j,k} \ldots \vec{y}) \text{ over all tuples of terms } \vec{t} \tag{1.2}$$

is not a theorem scheme. We wish to show that this scheme fails in the term model. So let $(\forall_\infty \vec{y})\phi(t_i, t_{i,j}, t_{i,j,k} \ldots \vec{y})$ be one of the instances that is not a theorem. Its negation is

$$(\exists_\infty \vec{y})\phi(t_i, t_{i,j}, t_{i,j,k} \ldots \vec{y})$$

and we wish to show that this is true in the term model. But this can be done by the constructions of theorem 1.6.1 and remark 9.

See section 1.8 for a discussion of the correct generalisation of this to random/generic/countably categorical structures.

1.6.2 Set theory and the cofinite quantifier

Is naïve set theory consistent in this language?

Extensionality is changed by the disappearance of equality. It becomes the assertion that

$$\forall_\infty x \forall_\infty y \exists_\infty z z \in x \Delta y$$

An axiom of comprehension becomes an assertion that there are infinitely many sets that each have only finite symmetric difference with the extension of ϕ. Since—for any ϕ—any two of these approximants have only finite symmetric difference the fact that there are infinitely many of them contradicts son-of-extensionality.

Aha! This is true only if $x \cup \{y\}$ exists for all x and y. But for this we need equality to be in the language! If we remove equality from the language then things might make more sense. Extensionality becomes the assertion that two sets with the same members belong to the same things.

1.7 Applications to WQO theory

A WQO is a quasiorder with no infinite antichains and no infinite strictly descending sequences. The class of BQOs is the largest class of WQOs closed under various operations: in particular if $\leq$ is a BQO of a set X, then the relation $\leq^+$ between subsets of X defined by $X' \leq^+ X''$ iff $(\forall u \in X')(\exists v \in X'')(v \leq v)$ is also BQO. Cute (unpublished) fact: if $\leq$ is a BQO of X then $(\forall_\infty u \in X')(\exists_\infty v \in X'')(v \leq v)$ is also BQO!!

Take the first-order theory of a wellordering (the theory of total orders plus an induction scheme). Do a surface translation into $\mathcal{L}^\infty$. Is it now the case that every model of this new theory is also a model of the first-order theory of antisymmetrical WQOs?

Must check that antisymmetry$^\infty$ implies antisymmetry.

1.7.1 Axiomatising WQO theory using the cofinite quantifier

I've been wondering if one can write down a set of axioms for WQOs using only the cofinite quantifiers instead of the usual ones. By that i meant that all the first-order quantifiers should be new-style; higher-order quantifiers might have to be old-style, but i'm not sure. For example, asserting that all but finitely many subsets have infinitely many minimal elements ensures that they all do, as if there is even one subset lacking a minimal element then there are infinitely many lacking minimal elements. OTOH i've just noticed that having *infinitely* many minimal element is a complication we can do without.

There are subtleties: definitions and axioms that look equivalent under the old interpretations might not be equivalent under the new. A lurking issue is whether or not the new axioms put any constraint on the sizes of the equivalence classes.

x is minimal: $(\forall y)(y \leq x \to x \leq y)$;

Trouble is: saying there are minimal elements involves saying there are infinitely many minimal elements. So some equivalence classes will have to be infinite.

No infinite antichains: $(\forall x)(\forall y)(x \leq y \vee y \leq x)$

Put these two together to infer that every element is above a minimal element. It is certainly true of each of cofinitely many x that there are at most finitely many things they aren't comparable with. But there are infinitely many minimal elements, so each of these x is comparable with a minimal element. So it's above a minimal element. But if there is even one x with no minimal element below it there must be infinitely many, by DC.

For every $f : \mathbb{N} \to X$ there are infinitely many i s.t. there are infinitely many $j > i$ with $f(i) \leq f(j)$. This sounds like perfect subarray lemma, but it isn't, and we have to be very careful. However, what we can be sure is that the condition that for any $f : \mathbb{N} \to X$ $(\exists_\infty i)(\exists_\infty j)(f(i) \leq f(j))$ isn't any stronger than the WQO condition, since it is implied by the perfect subarray lemma.

Here's an idea. Set up a first-order theory of wellorderings. The only subtlety—which we should be able to solve in short order—is: can do we use $\leq$ or $<$? It will have axioms to say that $\leq$ is a total order. Let's use $\leq$ instead of $<$ so that we can express connexity without using $=$.

It will also have an axiom scheme of induction

$$(\forall x)((\forall y < x)(\phi(y)) \to \phi(x)) \to \phi(x)) \to (\forall z)(\phi(z)) \tag{1.1}$$

This becomes

$$(\forall_\infty x)((\forall_\infty y < x)(\phi(y)) \to \phi(x)) \to \phi(x)) \to (\forall_\infty z)(\phi(z)) \tag{1.2}$$

Is this true of WQOs? Formula 1.2 is a conditional, as is formula 1.1. The conclusion of formula 1.2 is weaker than the conclusion of formula 1.1. However the premiss is weaker too, since it is a conditional whose premiss is weaker than the premiss of the corresponding part of formula 1.1.

Suppose, for all but finitely many x, that $\phi(x)$ follows as long as cofinitely many things below x are ϕ. Can there be infinitely many things that are not ϕ? If cofinitely many of your $<$-predecessors are ϕ you'd better be ϕ too when the music stops—unless you have a special pass, and there are only finitely many passes going. So let off with a caution the people who have special passes, and corral the rest: consider the set B of those counterexamples that have infinitely many counterexamples below them. We want B to be empty. If B is nonempty then every member of B not only has infinitely many counterexamples below it, but actually has infinitely many members of B below it. Note that B has no minimal element, contradicting wellfoundedness of $<$.

> A: if $\leq$ is WQO then $(\forall_\infty x)(\forall_\infty y)(y \leq x)$ is false but $(\forall_\infty y)(\forall_\infty x)(y \leq x)$ is true

even tho' $\forall_\infty$ is monotone.

At least i think this is true: it's certainly true for $\leq_{\mathbb{N}}$.

It's difficult not to smell a rat: $<_{\mathbb{N}}$ is a very asymmetrical relation. Is there a converse? That is to say, if we have a relation R such that $(\forall_\infty x)(\forall_\infty y)(R(x,y))$ but not $(\forall_\infty y)(\forall_\infty x)(R(y,x))$ then must R be a wellorder of transfinite length or at least give rise to one? This is a fruitful idea: One thing is clear from the outset, namely that any relation that satifies $(\forall_\infty x)(\forall_\infty y)(R(x,y))$ is at least a "well relation" in the sense of Marcone: If $\langle x_i : i \in \mathbb{N}\rangle$ is an infinite sequence of elements in the domain of R then there must be $i < j$ with $R(x_i, x_j)$. However it needs refinement, as the following counterexample shows. Consider the countable binary structure $\langle X, R\rangle = \{a_i : i \in \mathbb{N}\} \cup \{b_{i,j} : i,j \in \mathbb{N}\}$ and R = complement of $\{\langle a_i, b_{i,j}\rangle : i,j \in \mathbb{N}\}$. then $(\forall_\infty x)(\forall_\infty y)(R(x,y))$ is true but $(\forall_\infty y)(\forall_\infty x)(R(y,x))$ is false. However, R is not transitive. If we have a *transitive* relation R such that $(\forall_\infty x)(\forall_\infty y)(R(x,y))$ but not $(\forall_\infty y)(\forall_\infty x)(R(y,x))$ then we do indeed have something nice.

First let us notice that we can assume wlog that this R is reflexive, and so we can take it to be a quasiorder and write it with '$\leq$' thus $(\forall_\infty x)(\forall_\infty y)(x \leq y)$ but not $(\forall_\infty y)(\forall_\infty x)(y \leq x)$. This R is in fact not only a quasiorder but a WQO. We will show that it is a WQO of infinite rank, which is as good a converse to A as could be reasonably hoped for.

The intersection of a quasiorder and its converse is an equivalence relation, and we need to think about the cardinalities of the equivalence classes. Can any of these equivalence classes be infinite? Well, if there is an infinite equivalence it must be the top element of the quotient, beco's any element in an equivalence class $\not\leq$ it would not be able to see cofinitely many things.

Now if we had an element that was invisible to infinitely many things (as implied by the second half of A). How many equivalence classes do these infinitely many things belong to? By the above argument, these equivalence classes must each be finite, and so there must be infinitely many of them. But if a WQO has infinitely many equivalence classes it must be of infinite rank.

I'm sure there is more one can say than this ....

Given a preorder one can cook up a quantifier in one of two ways. There is the CLUB quantifier. This actually isn't very satisfactory beco's of—for example—the 2-ladder: domain is $\{a_i : i \in \mathbb{N}\} \cup \{b_i : i \in \mathbb{N}\}$ where each thing $\leq$ anything with greater subscript. Then the as and the bs are disjoint clubsets. This means that the club quantifier isn't a filter, which mucks things up. The other quantifier is the sufficiently large quantifier.

1.7.2 The cofinite quasiorders on the power set

Recall from the BQO notes the following:
If $\langle X, \leq \rangle$ is a quasiorder, define $\leq^{\aleph_0}$ on $\mathcal{P}(X)$ by
$X_1 \leq^{\aleph_0} X_2$ iff $(\forall_\infty x_1 \in X_1)(\exists_\infty x_2 \in X_2)(x_1 \leq x_2)$.
Then

Theorem 12. *If $\langle X, \leq \rangle$ is a BQO then so is $\langle \mathcal{P}(X), \leq^{\aleph_0} \rangle$.*

and
If we let $\leq_1$ be that relation $\leq^{\aleph_0}$ we have just seen, and for $n \geq 1$ we ordain

Definition 13. $Y \leq_{n+1} X$ *iff* $(\forall_\infty y \in Y)(\exists_\infty x \in X)(y \leq x \wedge (Y \setminus \{y\} \leq_{n-1} X \setminus \{x\})$

then

Theorem 14. *$\langle \mathcal{P}(X), \leq_n \rangle$ is BQO as long as $\langle X, \leq \rangle$ is.*

1.8 Conversation with John Truss

Some random logical thoughts:

Let G be the canonical random graph, and suppose G is a subgraph of another graph G'. Then every $\forall^*\exists^*$ sentence true in G' is true in G. But the same method will show that every ${\forall_\infty}^*{\exists_\infty}^*$ sentence true in G' is true in G. The thought is that all the logic-y sorts of things one can prove about G hold even when one reinterprets the quantifiers as $\forall_\infty$ and $\exists_\infty$.

Ask John Truss!

The canonical random graph has the finite model property: everything true in it has a finite model. Probably a consequence of the zero-one law. The direct limit of the family of finite triangle-free graphs is not known to have the same property. But John says it's beco's it's random rather than becos it's generic.

- Generic = done by Fraïssé 's method. A direct-limit of a suitable family K of finite structures. K must be closed under iso and substrux, have $\aleph_0$ isomorphism types, must have joint embedding and amalgamation. Amalgamation is like JEP but you have to take into account the intersection of the two strux.

- Random = R is the random structure associated with a class K if for every expression ϕ of the relevant language $R \models \phi$ iff the proportion of those K-structures of size n that are ϕ tends to 1 as $n \to \infty$.

Canonical random graph is both. Generic triangle-free graph is not the same as the random triangle-free graph, which is the same as the random bipartite graph.

Generic structures are pretty obviously homogeneous.

So the moral seems to be that it is *random* structures that we should be thinking of using the cofinite logic for. Is it the case that every finite wombat has a nice embedding into the countable random wombat?

The almost automorphism group of the countable random graph consists of things that preserve formulæ of $\mathcal{L}^\infty$.

A **near automorphism** of a structure is an isomorphism between two cofinite subsets, identified up to finite symmetric difference.

John says that games are useful, despite what i say: he says try proving that the concatenation of two copies of Z is elementarily equivalent as an ordered set to one copy—without using Ehrenfeucht-Fraïssé games.

How much of the apparatus for discussing superatomicity in Boolean algebras can be done here?

1.9 Slides

The language

Two quantifiers $\forall_\infty$ and $\exists_\infty$ which mean "for all but finitely many..." and "there are infinitely many..."

Prenex Normal Form

$$\exists_\infty xA \longleftrightarrow \exists_\infty yA[y/x]$$

(renaming of bound variables)

$$(\exists_\infty x)(A \wedge B) \longleftrightarrow (\exists_\infty xA) \wedge B$$

('x' not free in B)

$$\exists_\infty x(A \vee B) \longleftrightarrow (\exists_\infty xA) \vee B$$

('x' not free in B)

$$\exists_\infty x(A \vee B) \longleftrightarrow (\exists_\infty xA \vee \exists_\infty xB$$

$$\forall_\infty x(A \wedge B) \longleftrightarrow (\exists_\infty xA \wedge \exists_\infty xB$$

Elimination of equality

Torkel's axiomatisation

1) All instances of ordinary propositional axioms and identity axioms,
2) the further axioms (for every A and B)

1. $(\exists_\infty x)A \to A$ if x is not free in A;
2. $(\exists_\infty x)A(x) \to (\exists_\infty y)A(y)$;
3. $\neg(\exists_\infty x)(x = s)$;
4. $(\exists_\infty x)(A \vee B) \to (\exists_\infty x)A \vee (\exists_\infty x)B$;
5. $(\exists_\infty x)A(x) \to (\exists_\infty x)(\exists_\infty y)(A(x) \wedge A(y) \wedge \neg x = y)$.

We have the rule: if $A \to B$ has been proved, conclude $(\exists_\infty x)A \to (\exists_\infty x)B$". We need three little lemmas

Lemma 2 (The constant elimination property) If $G(c) \vdash A(c)$ where c is a constant, $G(x) \vdash A(x)$

(The constant elimination property.)

Lemma 15. *For every m, every G, every A and B: If $G \vdash A(x_1) \wedge \ldots \wedge A(x_m)$ and all the x_i are different $\to B(x_1) \vee \ldots \vee B(x_m)$ and $G \vdash \exists_\infty x A(x)$ then $G \vdash \exists_\infty x B(x)$.*

Lemma 16. *For every n, every $A_1, \ldots A_n$ and every B:*
If

$$G \vdash A_1(x_1) \wedge \ldots \wedge A_n(x_n) \to B(x_1) \vee \ldots \vee B(x_n)$$

then

$$G \vdash \exists_\infty x A_1(x) \wedge \ldots \wedge \exists_\infty x A_n(x) \to \exists_\infty x B(x)$$

Chapter 2

Branching Quantifiers

Branching quantifier formulæ look like

$$\begin{pmatrix} \forall x \exists y \\ \forall z \exists w \end{pmatrix} (A(x,y,z,w)) \tag{2.1}$$

The Hintikka game treatment of this makes `true` and `false` into *teams* of two players. (This is to do with the width of the antichains of quantifiers, which in this case is 2!) Team `false` starts: each member instantiates one of the two universal quantifiers. Then team `true` chooses witnesses for y and w *without collaborating*. That is to say, the member of `true` who chooses a witness for 'y' is told only about `false`'s choice for 'x', and the member of `true` who chooses a witness for 'z' is told only about `false`'s choice for 'w'. Members of team `true` are not allowed to impart information to each other. The game then proceeds as an ordinary Hintikka game, with the usual rules for termination. [1]

The easiest way to present the family of approximants is to exhibit the first two or so. The first one is

$$\forall x \forall z \exists y \exists w \; A(x,y,z,w)$$

The second is

$$(\forall x_1 \forall z_1 \exists y_1 \exists w_1)(\forall x_2 \forall z_2 \exists y_2 \exists w_2) \bigwedge \begin{pmatrix} A(x_1, y_1, z_1, w_1) \\ A(x_2, y_2, z_2, w_2) \\ x_1 = x_2 \to y_1 = y_2 \\ z_1 = z_2 \to w_1 = w_2 \end{pmatrix} \tag{ϕ_2}$$

The second approximant says that if `false` picks the same x (resp. z) for a second time, then `true` must reply with the same y (resp. w) as used the first time. If we constrain `true`'s play in the obvious way, by saying that if she has

[1] Of course the general branching-quantifier prefix is not a set of rows, but an arbitrary partial order, so that the members of the two teams are not indexed by *rows* as if they were rugby forwards but rather by maximal chains through the poset.

ever replied with a to a challenge b she must always reply with a to any future challenges with b we arrive at a family

$$(\forall x_1 \forall z_1 \exists y_1 \exists w_1) \ldots (\forall x_n \forall z_n \exists y_n \exists w_n) \begin{pmatrix} \bigwedge_{i \leq n} A(x_i, y_i, z_i, w_i) \\ \bigwedge_{i < j \leq n} (x_i = x_j \to y_i = y_j) \\ \bigwedge_{i < j \leq n} (z_i = z_j \to w_i = w_j) \end{pmatrix} \qquad (\phi_n)$$

Let's call this formula ϕ_n. The family of approximants indexed by $\mathbb{N}$ that we elaborate was shown by Barwise to have the following features.

- $\forall n \in \mathbb{N}\ \phi \to \phi_n$ is valid. The easy way to see this is to spot that it becomes easier for team `True` to win if they are told more of `False`'s choices for the universal quantifiers, so if they can win the branching quantifier formula they can certainly win the approximants.
- $\forall n \in \mathbb{N}\ \phi_{n+1} \to \phi_n$ is valid
- Any first-order theory consistent with all the ϕ_n is consistent with ϕ.

For the Barwise-Hintikka team game to give a correct semantics for formulæ with branching quantifiers it is necessary for team `true` to be prevented from cheating. If they cheat they might well end up winning plays that they shouldn't.

The secret to making sense of the higher approximants is to realise that they are interesting only in the case where `false` checks `true`'s honesty by ensuring that at each position at least *one* of the things they pick is the same as one they picked earlier.

true have a strategy to win the truth game for the branching-quantifier formula ϕ without cheating as long as the formula is true. They have a strategy to win the truth game for ϕ_n if they are allowed to cheat for fewer than n moves. (in the truth game for ϕ_n **true** can see all **false**'s moves for the earlier universal quantifiers)

One good way to think of ϕ_n is through the following game. At stage n player `false` picks x_n and z_n, and player `true` replies with y_n and w_n. Unless $A(x, y, z, w)$, `false` wins at once. If $x_n = x_j$ for some earlier j then y_n must equal y_j or `true` loses. Similarly if $z_n = z_j$ for some earlier j then w_n must equal w_j or `true` loses. If `true` survives these hazards they procede to stage $n + 1$.

Now we find that the Hintikka game for ϕ_n is the truncation of this game to n moves, where player `true` wins the truncation if she is still in the game after n moves.

If ϕ is a formula that says that two structures are isomorphic it takes only a further slight rearrangement to obtain the Ehrenfeucht-Fraïssé game that describes the same thing.

Let's Skolemize ϕ_n. y_i depends on $x_1 \ldots x_i$ and $z_0 \ldots z_{i-1}$ and w_i depends on $x_1 \ldots x_i$ and $z_1 \ldots z_i$. Using $f_i(x_1 \ldots x_i, z_0 \ldots z_{i-1})$ for y_i and $g_i(x_1 \ldots x_i, z_1 \ldots z_i)$ for w_i we obtain

$$(\forall x_1 \ldots x_n)(\forall z_0 \ldots z_{n-1}) \begin{cases} \bigwedge\limits_{i \leq n} A(x_i, f_i(x_1 \ldots x_i, z_0 \ldots z_{i-1}), z_i, g_i(x_1 \ldots x_i, z_1 \ldots z_i)) \\ \bigwedge\limits_{i<j \leq n} (x_i = x_j \to f_i(x_1 \ldots x_i, z_0 \ldots z_{i-1}) = f_j(x_1 \ldots x_j, z_0 \ldots z_{j-1})) \\ \bigwedge\limits_{i<j \leq n} (z_i = z_j \to g_i(x_1 \ldots x_i, z_0 \ldots z_i) = g_j(x_1 \ldots x_j, z_0 \ldots z_j)) \end{cases}$$

(ϕ_n: Skolemized)

Team `true` have the Skolem functions f_i and g_i to play with. The second and third row conjunctions say that the values of f_i depend only on its x inputs and the values of g_i depend only on the z inputs. If `true` cheat they do not respect this constraint and by the rules of the Hintikka game they lose.

Barwise says: ϕ_n is equivalent to the assertion that **true** can cheat so successfully that it can remain undetected for n repeats. I don't understand this yet

ϕ could be any branching-quantifier formula of course, but if we want to connect this to Ehrenfeucht games we will need to consider the particular case of the branching-quantifier formula that says there is an isomorphism between two structures. (It's standard that this can be done with branching quantifiers.) To keep the formulæ readable i shall consider even a special case of that, namely the formula that says that there is an antimorphism of the universe which is an involution.[2] I take this example for a number of reasons. First, if we merely want to say there is a bijection betwen A and B then we will end up with a Henkin prefix of width 4 which is horrible. Secondly if we take other, apparently more natural examples of things with bijections between them (two countable total orders) then we have to embed them in some larger structure to be able to use this rhetoric. This involves making the disjoint union of two languages, which is a chore, and it's much easier to talk about a bijection between the universe and its dual.

This is

$$\begin{pmatrix} \forall y_1 \exists x_1 \\ \forall y_2 \exists x_2 \end{pmatrix} \bigwedge \begin{pmatrix} x_1 = x_2 \longleftrightarrow y_1 = y_2 \\ x_1 = y_2 \longleftrightarrow y_1 = x_2 \\ y_1 \in y_2 \longleftrightarrow x_1 \notin x_2 \end{pmatrix} \tag{2.2}$$

Take a minute or so to look at this formula. It's a bit of a bugger, but better than most. The clause $x_1 = y_2 \longleftrightarrow y_1 = x_2$ ensures that the function whose existence is being asserted is actually an involution. Perhaps we should abbreviate $y_1 \neq x_1 \wedge y_2 \neq x_2 \wedge (y_1 = y_2 \longleftrightarrow x_1 = x_2) \wedge (x_1 = y_2 \longleftrightarrow x_2 = y_1)$ to $\Box(y_1, x_1, y_2, x_2)$ so the formula then becomes

$$\begin{pmatrix} \forall y_1 \exists x_1 \\ \forall y_2 \exists x_2 \end{pmatrix} ((\Box(y_1, x_1, y_2, x_2) \wedge (y_1 \in y_2 \longleftrightarrow x_1 \notin x_2))) \tag{2.3}$$

[2] I'm not sure that this has the same consistency strength as the existence of an unvarnished antimorphism but it's only an illustration anyway

The first approximant to this is

$$(\forall y_1 y_2)(\exists x_1 x_2)(\Box(y_1, x_1, y_2, x_2) \wedge (y_1 \in y_2 \longleftrightarrow x_1 \notin x_2))$$

Then it is not hard to persuade oneself that this scheme is actually going to be equivalent to the scheme.

$$(\forall y_1)(\exists x_1) \ldots (\forall y_n)(\exists x_n)(\langle \vec{y}, \in, = \rangle \simeq \langle \vec{x}, \notin, = \rangle)$$

Of course in the more general context where one is considering approximations to the branching quantifier formula that says that one structure is isomorphic to another the quantifier prefix looks exactly the same, but the matrix is $Y \simeq X$ or something like that.

The Hintikka game for this formula is simply the nth Ehrenfeucht game.

Any first-order theory consistent with all the ϕ_n is consistent with ϕ. This implies than any first-order consequence of ϕ is a consequence of a ϕ_n.

Let's be more specific, and work on a live example.

Consider the language of graph theory (it has a binary relation $E(x, y)$ which says that there is an edge joining x and y) with k extra constant symbols $v_1 \ldots v_k$. (These are k designated vertices each encoding a different colour). Now look at the expression

$$\begin{pmatrix} \forall x_1 \exists y_1 \\ \forall x_2 \exists y_2 \end{pmatrix} \bigwedge \begin{cases} x_1 = x_2 \to y_1 = y_2 \\ E(x_1, x_2) \to y_1 \neq y_2 \\ \bigvee\limits_{i \leq k} (y_1 = v_i) \\ \bigvee\limits_{i \leq k} (y_2 = v_i) \end{cases} \tag{2.4}$$

This says that there is a map defined on all vertices, taking values in a k-sized set, which sends vertices connected by an edge to distinct members of the set. It can be true only in a graph that is k-colourable. It is well-known that for $k > 2$, being k-colourable is an NP-complete property, and is also Σ^2_1. How? "There are k disjoint subsets of the set of vertices s.t. ...".

Walkoe [1970] proves that *any* Σ^2_1 property can be captured by branching quantifiers if one uses a trick like this. Actually, he doesn't prove it, but takes the reader through an example complicated enough for the reader to get the idea how to do it in general (*op. cit.* p 541.) In fact most of the work for this proof has been done for us earlier in this section by Ehrenfeucht: the idea is to use a clause like $(x_1 = x_2 \to y_1 = y_2)$ to code an injection.

THEOREM 17. *(Fagin-Walkoe) the following are equivalent*

1. Φ *expressible with branching quantifiers*

2. $\Phi \in \Sigma^2_1$

3. Φ *restricted to finite sets is in* NP

and all of these imply
4. Φ preserved by ultraproducts.
$1 \rightarrow 2$, $2 \rightarrow 4$ are easy. Complete this proof

2.1 A factoid concerning width of branching quantifier prefixes

I start with a simple illustration. Consider the formula

$$(\forall x)(\exists y)(\forall u)(\exists v)\phi(x, y, u, v) \tag{2.1}$$

I claim that this is equivalent to the branching-quantifier formula

$$\begin{pmatrix} \forall x \exists y \\ \forall x' \forall u \exists v \end{pmatrix} (x = x' \rightarrow \phi(x, y, u, v)) \tag{2.2}$$

To see this, simply Skolemize formula (2.1) to obtain

$$(\forall x)(\forall u)\phi(x, f_1(x), u, f_2(x, u)) \tag{2.3}$$

So if formula (2.1) is true, in the Hintikka game for formula (2.1) player `true` has the two Skolem functions f_1 and f_2. In the Barwise-Hintikka game for formula (2.2) player `true`'s team has two members `Top-row` and `Bottom-row`. Obviously `Top-row` should use f_1 and `Bottom-row` should use f_2 on x' and u.

What can `false`'s two team members—`Top-row` and `Bottom-row`—do? Unless they make the same choice for x and x' their team loses. The only way they can win is to ensure that $x = x'$ and if they do, `true`'s strategy wins.

For the other direction we notice that for the two players of team `true` to have a strategy to win the Barwise-Hintikka game for formula (2.2) they cannot rely on team `false` making a pig's ear of things by picking $x \neq x'$, so they must have the two functions f_1 and f_2. But then `true` can win formula (2.1).

In this case we were dealing with a $\forall^1\exists^1\forall^1\exists^1$ formula. There is a trivial generalisation of this case to $\forall^*\exists^*\forall^*\exists^*$ formulæ. It's a bit messy:

$$(\forall \vec{x})(\exists \vec{y})(\forall \vec{u})(\exists \vec{v})\phi(\vec{x}, \vec{y}, \vec{u}, \vec{v}) \tag{2.4}$$

turns out to be equivalent to

$$\begin{pmatrix} \forall \vec{x} \exists \vec{y} \\ \forall \vec{x}' \forall \vec{u} \exists \vec{v} \end{pmatrix} \left(\bigwedge x_i = x_i{}' \rightarrow \phi(\vec{x}, \vec{y}, \vec{u}, \vec{v})\right) \tag{2.5}$$

...and for the same reasons. The feature not to be lost sight of is that we end up with a branching-quantifier formula whose prefix is of width two as before.

The interesting generalisation concerns the case where we have more pairs of quantifiers (or blocks of quantifiers) than the two we had in the case just considered. The best way to illustrate the general strategy is probably to consider a formula with *four* pairs of quantifiers (which is to say a $\forall_8$ formula):

$$(\forall x)(\exists y)(\forall u)(\exists v)(\forall a)(\exists b)(\forall c)(\exists d)\phi(x,y,u,v,a,b,c,d) \tag{2.6}$$

This is logically equivalent to the branching-quantifier formula:

$$\begin{pmatrix} \forall x \exists y \\ \forall x' \forall u \exists v \\ \forall x'' \forall u' \forall a \exists b \\ \forall x''' \forall u'' \forall a' \forall c \exists d \end{pmatrix} \left(\bigwedge \left\{ \begin{array}{l} x = x' = x'' = x''' \\ u = u' = u'' \\ a = a' \end{array} \right\} \rightarrow \phi(x,y,u,v,a,b,c,d) \right) \tag{2.7}$$

By thinking about the Skolemized version of formula (2.6) we can see that player `true` will have a winning strategy for the Hintikka game for formula (2.6) as long as she has the the four obvious Skolem functions, $f_1 \ldots f_4$ of 1, 2, 3 and 4 arguments respectively, which are the analogues of f_1 and f_2 earlier. But if she *has* got these four functions, then they can be doled out in the obvious way to the four players in `true`'s team for the Barwise-Hintikka game for formula (2.7): `first-row` uses f_1, `second-row` uses f_2, `third-row` uses f_3 and `fourth-row` uses f_4. Thus equipped, team `true` will win.

This illustrates the following

THEOREM 18. *Every $\forall_{2n}$ formula is equivalent to a branching-quantifier formula of width n where each branch of the prefix is $\forall^*\exists^*$.*

Proof:

Although we approached theorem 18 through games, we can actually prove theorem 18 without talking about games at all. We illustrate how, by transforming formula (2.2) into formula (2.1) using only reversible operations that preserve logical equivalence.

We start with formula (2.2):

$$\begin{pmatrix} \forall x \exists y \\ \forall x' \forall u \exists v \end{pmatrix} (x = x' \rightarrow \phi(x,y,u,v))$$

...Skolemize it $((f_1(x)/y$ and $f_2(x',u)/v)$ to obtain

$$\begin{pmatrix} \forall x \\ \forall x' \forall u \end{pmatrix} (x = x' \rightarrow \phi(x, f_1(x), u, f_2(x',u))) \tag{2.8}$$

Clearly we are allowed to squash flat any branching quantifier prefix that consists of only one kind of quantifier (as long as its a self-commuting quantifier like $\forall$ or $\exists$) so formula (2.8) is equivalent to

$$\forall x \forall x' \forall u (x = x' \rightarrow \phi(x, f_1(x), u, f_2(x',u)) \tag{2.9}$$

Next we exploit the logical equivalence of $(\forall x_1)(\forall x_2)(x_1 = x_2 \rightarrow \phi(x_1))$ with $(\forall x)\phi(x)$ to conclude that formula (2.9) is equivalent to

$$(\forall x)(\forall u)(\phi(x, f_1(x), u, f_2(x,u)) \tag{2.10}$$

And formula (2.10) is the Skolemized version of

$$(\forall x)(\exists y)(\forall u)(\exists v)\phi(x,y,u,v)$$

which of course is formula (2.1). ■

So my first question is: is this result new?

2.2 Skolemisation

I see how to do $\exists_\infty F(x)$ in the branching quantifier language. You invent a constant and say that it has F, and then say there is a bijection between the extension of F and the extension of $F \setminus \{a\}$:

$$\left(\begin{array}{c}\forall x\exists y\\ \forall u\exists v\end{array}\right)\bigwedge\left\{\begin{array}{l}x=u\longleftrightarrow y=v\\ (F(x)\to F(y))\wedge(F(u)\to F(v))\\ y\neq a\wedge v\neq a\\ F(a)\end{array}\right)$$

With a little bit of work i can get rid of the constant:

$$\left(\begin{array}{c}\forall a\forall x\exists y\\ \forall b\forall u\exists v\end{array}\right)(a=b\wedge F(a))\to\bigwedge\left\{\begin{array}{l}x=u\longleftrightarrow y=v\\ (F(x)\to F(y))\wedge(F(u)\to F(v))\\ y\neq a\wedge v\neq a\\ F(a)\end{array}\right\}$$

It would be nice to do $\forall_\infty F(x)$ in the branching quantifier language too.

Notice that Skolemisation preserves satisfiability but not validity. $(\forall x)(\exists y)(x = y)$ is valid but its skolemisation (which is $(\forall x)(x = f(x))$) is not valid! Indeed

Remark 19. *Except in trivial cases, the skolemised version of a formula is never valid*

Proof:

Suppose ϕ' is the skolemisation of ϕ. Then each *new* function letter (arising from Skolemisation) in ϕ' appears only in connection with one variable. If the function letter 'f' arose by skolemising the variable 'y' which is in the scope of 'x' then, in ϕ', 'f' only ever appears succeded by 'x'. If we have a proof of ϕ' we can replace all occurrences of '$f(x)$' in it by a new constant, and then use UG on that constant. So if ϕ' were valid, then the formula obtained from ϕ by replacing all existential quantifiers by universal quantifers would be valid too. ■

When we Skolemize a formula, we replace existentially bound variables by terms like '$f(\vec{x})$'. However, although different Skolem function letters can appear with the same argument (if there are two adjacent existential quantifiers in the prefix) *each function letter can only ever appear with the one (tuple of) argument(s)!* (The old existentially bound variables are in 1-1 correspondence with *function letters*, not with *terms-composed-of-those-letters*). This will prevent us from saying that f is injective for example.

It's worth thinking about how branching quantifiers subvert this. Skolemize a branching-quantifier formula; what happens?

Take as an illustration the formula that says that a graph is k-colourable:

$$\begin{pmatrix} \forall x_1 \exists y_1 \\ \forall x_2 \exists y_2 \end{pmatrix} \bigwedge \begin{cases} x_1 = x_2 \to y_1 = y_2 \\ E(x_1, x_2) \to y_1 \neq y_2 \\ \bigvee_{i \leq k} (y_1 = v_i) \\ \bigvee_{i \leq k} (y_2 = v_i) \end{cases}$$

Let us Skolemize this, using 'f' for the first-row Skolem function and 'g' for the second row.

$$\begin{pmatrix} \forall x_1 \\ \forall x_2 \end{pmatrix} \bigwedge \begin{cases} x_1 = x_2 \to f(x_1) = g(x_2) \\ E(x_1, x_2) \to f(x_1) \neq g(x_2) \\ \bigvee_{i \leq k} (f(x_1) = v_i) \\ \bigvee_{i \leq k} (g(x_2) = v_i) \end{cases}$$

Clearly we can flatten this prefix as before, so this is equivalent to

$$(\forall x_1)(\forall x_2) \bigwedge \begin{cases} x_1 = x_2 \to f(x_1) = g(x_2) \\ E(x_1, x_2) \to f(x_1) \neq g(x_2) \\ \bigvee_{i \leq k} (f(x_1) = v_i) \\ \bigvee_{i \leq k} (g(x_2) = v_i) \end{cases}$$

The first line tells us that $f = g$ which simplifies things mightily. We get

$$(\forall x_1)(\forall x_2) \bigwedge \begin{cases} E(x_1, x_2) \to f(x_1) \neq f(x_2) \\ \bigvee_{i \leq k} f(x_1) = v_i \\ \bigvee_{i \leq k} f(x_2) = v_i \end{cases}$$

$$(\forall x_1)(\forall x_2)(E(x_1, x_2) \to f(x_1) \neq f(x_2). \wedge . \bigvee_{i \leq k} f(x_1) = v_i \wedge \bigvee_{i \leq k} f(x_2) = v_i)$$

But notice that the Skolem function f now appears with two *different* arguments—and therefore could never have arisen by Skolemizing a first-order formula! If you skolemise a first-order formula then you might end up with two skolem functions being fed the same list of arguments; what cannot happen is two distinct lists of arguments being fed to the one skolem functions. This happened because f was originally two distinct Skolem functions (arising from different prefixes), which we were able to compel to have the same graph by judicious use of equality. I can see no way of doing this without exploiting the presence of '=' in the language.

This raises the second question: Is branching quantifier logic without equality any stronger than LPC? I don't see how to prove it, even if it's true. Here is an obvious obstacle. Consider

$$\begin{array}{l}\forall y_1 \exists x_1 \\ \forall y_2 \exists x_2\end{array} (\phi(y_1, y_2, x_1, x_2)) \tag{2.1}$$

where ϕ contains no equalities. The problem is: if this is to be equivalent to a formula of LPC, which of the two following formulæ is it to be?

$$\forall y_1 \exists x_1 \forall y_2 \exists x_2 \phi(y_1, y_2, x_1, x_2) \tag{2.2}$$

$$\forall y_2 \exists x_2 \forall y_1 \exists x_1 \phi(y_1, y_2, x_1, x_2) \tag{2.3}$$

They are different, and any argument that 2.1 is equivalent to one of 2.2 and 2.3 is an argument that it's equivalent to the other! And please don't tell me that 2.2 and 2.3 are the equivalent if the language lacks equality, beco's i don't believe it.

Come to think of it it could perhaps be

$$\forall y_1 \forall y_2 \exists x_1 \exists x_2 \phi(y_1, y_2, x_1, x_2) \tag{2.4}$$

Another reason why i don't believe it is that the branching quantifier language with $\forall_\infty$ and $\exists_\infty$ (which lacks equations) can say things that the first-order linear version can't—because like quantifiers do not commute.

$$\begin{pmatrix}\forall_\infty y_1 \\ \forall_\infty y_2\end{pmatrix} \phi(y_1 y_2)$$

is not the same as

$$\forall_\infty y_1 \forall_\infty y_2 \phi(y_1 y_2)$$

nor the same as

$$\forall_\infty y_2 \forall_\infty y_1 \phi(y_1 y_2)$$

$$\left(\begin{array}{c}\forall_\infty y_1 \\ \forall_\infty y_2\end{array}\right) \phi(y_1 y_2)$$

is really

$$\left(\begin{array}{c}(\exists Y_1 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_1 \in Y_1) \\ (\exists Y_2 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_2 \in Y_2)\end{array}\right) \phi(y_1 y_2)$$

Skolemizing gives us

$$\left(\begin{array}{c}(Y_1 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (\forall y_1 \in Y_1) \\ (Y_2 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (\forall y_2 \in Y_2)\end{array}\right) \phi(y_1 y_2)$$

$$(Y_1 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (Y_2 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (\forall y_1 \in Y_1)(\forall y_2 \in Y_2)\phi(y_1 y_2)$$

which is of course the same as

$$(Y_1 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (Y_2 \in \mathcal{P}_{cof}(\mathcal{D})) \wedge (\forall y_2 \in Y_2)(\forall y_1 \in Y_1)\phi(y_1 y_2)$$

$$(\forall y_2 \in Y_2)(\forall y_1 \in Y_1)\phi(y_1 y_2)$$

More stuff to fit in

We should really consider more complicated quantifiers to get the full flavour, since the Hintikka Game treatment of quantifiers is really an ellipsis.

Let's think about

$$\left(\begin{array}{c} \forall_\infty y_1 \\ \forall_\infty y_2 \end{array} \right) \phi(y_1, y_2)$$

This is really

$$\left(\begin{array}{c} (\exists Y_1 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_1 \in Y_1) \\ (\exists Y_2 \in \mathcal{P}_{cof}(\mathcal{D}))(\forall y_2 \in Y_2) \end{array} \right) \phi(y_1 y_2)$$

Skolemizing and flattening will give us

$$(\exists Y_1 Y_2)(\forall y_1 \in Y_1)(\forall y_2 \in Y_2)\phi(y_1 y_2)$$

where the Ys are variables over cofinite sets. This is emphatically *not* the same as either

$$(\forall_\infty y_1)(\forall_\infty y_2)\phi(y_1, y_2)$$

or

$$(\forall_\infty y_2)(\forall_\infty y_1)\phi(y_1, y_2)$$

...which are of course distinct from each other.

****

Let $A(x, y)$ be the payoff of x played against y. Consider the formula

$$\left(\begin{array}{c} \forall y \\ \exists x \end{array} \right) (A(x, y) \geq k) \tag{2.5}$$

and the truth-game played over it. It is a game of imperfect information beco's it has simultaneous moves. The result is that neither player has a winning strategy in the truth-game, and there is no result to the effect that the fmla is true iff **true** has a winning strategy—at least not if we are to believe bivalence!

There are in fact two games (tho' this is not a superposition game!), namely the two truth games for $(\forall y)(\exists x)(A(x, y) \geq k)$ and $(\exists x)(\forall y)(A(x, y) \geq k)$

$$(\exists r)(\forall c)(\forall r')(\exists c')(A(r, c) \geq A(r', c'))$$

says there is a row whose minimum is greater than all other row-minima

$$(\exists c)(\forall r)(\forall c')(\exists r')(A(r, c) \leq A(r', c'))$$

says there is a column whose maximum is no greater than any other column maximum.

Skolemising and flattening gets us nowhere. Is there anything illuminating we can say?

****

Skolemise $(\forall_\infty x)(\exists_\infty y)\phi(x, y)$ to obtain

$(\forall_\infty x)(\forall n, m \in \mathbb{N})(f(x, m) \neq f(x, m) \wedge \phi(x, f(x, m) \wedge \phi(x, f(x, n))))$

If we don't like having two sorts of universal quantifiers in this formula we can replace the numerical universal quantifiers by $\forall_\infty$ numerical quantifiers. But then we have the old worry about order of quantifier

Chapter 3

Old notes from Pulman-Gordon project

Uncountably many

More than half

Of measure one

use of measure quantifier makes theory of D decidable.

The quantifiers i have been talking about so far bind one variable at a time. Now life can actually be slightly more complicated, for we can consider quantifiers that bind MORE than one variable at a time. We can think of these quantifiers as things of type $(V^n \to \mathit{bool}) \to \mathit{bool}$. The simplest example of this is also the earliest—Carol Karp, infinitary universal quantifier. Shows us an example of something we can do with funny quantifiers but not with ordinary ones (define wellorder). It is hardly surprising that we can say nasty 2^{nd} order things with infinitary quantifiers, beco's they enable us to quantify directly over infinite subsets. What is more surprising is that we can get similar strength from branching quantifiers. These were first considered by Henkin [6] in the late '50's, and Ehrenfeucht proved that a logic with branching quantifiers is richer than LPC, in that we can define a predicate "infinite" in it. See Walkoe, [11] JSL 1970 or thereabouts. The first application to natural language is due i think to Hintikka—see Barwise [1].

Most

One quantifier to look at is Most, or M for short. Let's bear in mind the axioms

$$(Mu)(My)Fuy \longleftrightarrow (My)(Mu)Fuy$$

which is self-explanatory, and

$$(Mu)(F(u)) \to (My)(Mz)(\neg(y = z)/\ F(z))$$

which says that if $\{x : F\}$ is large, then so is $\{x : F\} \setminus \{w\}$ for "most" w (tho' obviously not for all w!)

In fact we have to treat "most" as part of a determiner rather than a quantifier beco's if we read "most P are Q" as $(Mu)(Pu \to Qu)$ then if most things aren't P then $(Mu)(Pu \to Qu)$ is vacuously true, and this leaves a nasty classical taste in the mouth. Steve Pullman says the way to look at this is to say that if we want to represent "all pigs fly" and "some pigs fly" as being statements about everything, then we can find a connective to fit in the middle ($\to$ or $\wedge$) but with "most" we can't. Counterexample in Barwise and Cooper somewhere.

Thus the axioms should really be

$$(MQu)(MQv)F \longleftrightarrow (MQv)(MQu)F$$

$$(MQu)(F(u)) \to (MQv)(MQz)(\neg(v = z) \wedge F(z))$$

(which is a more general version of $\forall x.MQy\lambda y.\neg(x = y)$—which is obvious, for any x, most Q are distinct from x! We cannot have

$(MQu)(F(u)) \to \forall v.(MQz)(\neg(v = z) \wedge F(z))$

We should consider the possibility of axiomatising, say, well-quasi-orders or κ-like orderings by means of Mosts

If we are to implement reasoning about "most" widgets it is natural to try to do it in the same way as we implement LPC, namely by elimination of quantifiers in favour of a sophisticated term-making device. In LPC we use ϵ-terms for $\exists$ and just drop $\forall$ altogether. We can't just drop (Most frog) quantifiers: consider how it might go. I have a (PROLOG-style) goal: $p \to q$. If i have a premiss that looks like $r \to p'$ I am in business if it is closed or is has a secret $\forall$ outside it, for then i can unify p and p'. But suppose it is a most-frogs quantifier? Then i can only unify the two p s if i know that the variable inside the occurrence of p can denote a frog. Even if i can unify them what can i do with it? With most frogs i can infer q' from r', so even if i could show r' i wouldn't have a proof of q'. Could i represent that as a proof of some modalised version of q'? But in any case where do i keep safely the knowledge that the variable can denote a frog so it is handy when i want it?

So we can't simply strip off (Most frog) in the way we can strip off $\forall$. It is probably a point worth making that it is rather a special feature of $\forall$ that we can do this. After all, if all we knew about $\exists$ and $\forall$ was that they are duals, one could expect no better than to have to hack out $(\forall x)\Phi(x)$ in favour of $\Phi((\epsilon x)(\neg\Phi(x)))$. We should think of just being able to drop $\forall$ as a piece of good luck.

So all we can expect here is to be able to invent a pseudo-ϵ term for the dual quantifier to (Most frogs). Let's try. I am going to call them ρ-terms, because 'ρ' looks like p, which is short for 'plenty' and 'plenty' is arguably the dual of many: if i wish to contradict someone who says that most frogs are not green, i say "plenty of frogs are green".

ρ-terms

Under this system whenever we see a wff

plenty(frog(x))(green(x))

we invent a term

(ρx)(frog(x)) $\wedge$ (green(x))[1]

So that

M(frog(x))(green(x)) becomes

(frog((ρx)($\neg$frog(x)$\vee\neg$green(x))) $\rightarrow$ (green((ρx)($\neg$frog(x)$\vee\neg$green(x)))

Now since the generic ρ object is presumably at least a frog we can discharge the antecedent to get

$$(green((\rho x)(\neg frog(x) \vee \neg green(x)))$$

The trouble with ρ-terms is that they give us no way of deriving Dick's inference.

```
Most A are B  Most A are C
--------------------------
     some B are C
```

The modal operator μ

Suppose i have Most frog ϕ and frog fred. I cannot infer ϕ fred but can i invent a modal operator μ such that i can infer μ(ϕ fred)? μ means "most-likely"

What axioms do we have for it?

(Most frog x)(ϕx) $\rightarrow$ ($\forall$ frog x)(μ(ϕx))

contrapose and substitute $\neg\phi$ for ϕ

$\neg$($\forall$ frog x)(μ($\neg\phi x$)) $\rightarrow$ $\neg$(Most frog x)($\neg\phi x$)

($\exists$ frog x)($\neg\mu$($\neg\phi x$)) $\rightarrow$ $\neg$(Most frog x)($\neg\phi x$)

Can we draw any inferences from it? Yes! If we have a goal

$$p(y) \rightarrow q(y)$$

and a premiss $(Most\ (\lambda x)F(x))((\lambda z)p(z))$ we can unify y with z to get a new goal $F(y) \rightarrow \mu q(y)$.

[1] Or should this be $(\rho x)(frog(x)) \rightarrow (green(x))$?

Is it a normal modal logic?

$$\frac{F(a) \rightarrow G(b)}{\mu(F(a)) \rightarrow \mu(G(b))}$$

is this true? well, $\mu(F(a))$ is just (Most frogs$x$)$(F(x))\wedge$frog(a) and there is no reason to suppose that if we add $F(a)$ as a premiss we can get $\mu(G(b))$ so, no, it isn't normal. It does satisfy necessitation tho'.

How about

$$\frac{\mu(F(a) \rightarrow G(b))}{\mu(F(a)) \rightarrow \mu(G(b))}?$$

Well, $\mu(F(a) \rightarrow G(b))$ is ... well, that depends on what the μ applies to ... best forget all about the modal logic!

Can we use this μ to simulate all inferences to do with *mosts*? Pretty obviously not. In any case it is not hard to persuade ourselves that there is no modal operator μ which verifies the following biconditional. Unless μ burble implies burble the right-hand side could be true and the left-hand side false. Unfortunately the desired interpretation of μ makes the inference from μ burble to burble impossible.

$Most\ frog\ green \iff (\forall x)(frog(x) \rightarrow \mu green(x))$

The fact that the left-to right implication works is probably nearly good enough, for we are interested in deriving conclusions which do not contain "most".

Most frog green $\rightarrow$ green$((\epsilon x)(\mu(\neg$ frog$(x))))$

Now suppose

Most frogs are green

we conclude

$(\forall x)(\text{frog}(x) \rightarrow \mu\ (\text{green}(x)))$

frog $(\epsilon x)(\text{frog}(x) \wedge \neg(\mu(\text{green}(x)))) \rightarrow$

$\mu(\text{green}(\epsilon x)(\text{frog}(x) \wedge \neg(\mu(\text{green}(x)))))$

Now the antecedent is certainly true: the epsilon object is a frog, so we conclude

$\mu(\text{green}(\epsilon x)(\text{frog}(x) \wedge \neg(\mu(\text{green}(x)))))$

and the question then is: do we feel that this retains enough of the force of the original?

If most A are B is it the case that most $\neg A$ are $\neg B$?

If most A are B and x is $\neg B$ is it $\mu\neg A$?

Bibliography

[1] Jon Barwise Branching quantifiers in English, JPL vol 8 (1979) p 47-80.

[2] Barwise, K.J.J. Some applications of Henkin quantifiers. Israel J of Maths **25** 1976 pp 47-63.

[3] Bell and Slomson, Models and Ultraproducts, Ch. 13, "Generalized Quantifiers",

[4] Fagin R. Generalised first-order spectra and polynomial-time recognisable sets. *in* R. Karp ed., Complexity of Computation SIAM/AMS proc 7 43-73 AMS/SIAM [1974]

[5] Thomas Forster. Term models for a weak set theory with a universal set. *Journal of Symbolic Logic* v 52 (1987) pp 374-387. (Reprinted in Follesdal, ed: Philosophy of Quine, V. Logic, Modality and Philosophy of Mathematics).

[6] Leon Henkin "Some Remarks on Infinitely Long Formulas", Infinitistic Methods, Proceedings of the Symposium on Foundations of Mathematics, Warsaw, 1959.

[7] Hintikka K.J.J. [1973] Logic, Language Games and Information O.U.P. Clarendon Press 1973

[8] Hintikka K.J.J. [1979] "Is", semantical games and semantical relativity. JPL v 8 pp 433-468

[9] David Marker, Theodore A. Slaman. Decidability of the Natural Numbers with the Almost-All Quantifier arXiv:math/0602415v1 [math.LO]

[10] Monk, Mathematical Logic, Chapter 30, "Finitary Extensions [Cardinality Quantifiers]",

[11] Walkoe W.F. Finite partially-ordered quantification. Journal of Symbolic Logic v 35 1970 pp 535-555

[12] Yasuhara, Mitsuru, "Syntactical and semantics, properties of generalized quantifiers," JSL, vol. 31, (1966), pp. 617-632.