Xavier Caicedo

Hilbert's ε-Symbol in the Presence of Generalized Quantifiers

Introduction

The so called Hilbert's ε-symbol transforms a formula $\varphi(x)$ in a term $\varepsilon x\,\varphi(x)$ with the intended meaning: "some x such that $\varphi(x)$, if such x exists, an arbitrary individual, otherwise". We will asume it is extensionally determined in the sense that it satisfies the scheme:

A1. $\forall x(\varphi \leftrightarrow \psi) \rightarrow (\varepsilon x\,\varphi = \varepsilon x\,\psi)$.

Hilbert proposed originally:

A2. $\exists x\,\varphi(x) \rightarrow \varphi(\varepsilon x\,\varphi(x))$,

as the only axiom scheme governing ε, [9], [10], but he refers in [9] to a choice function as the intended interpretation of ε, which suggests scheme A1. This scheme appeared first explicitly in Ackermann [2]. Asser [1] showed later that the two schemes (plus a technical scheme and first order rules) form a complete system for first order logic enriched with the choice function interpretation of ε. Leisenring [11] gives a good account of this logic.

Although it has played an important role in proof theory (Hilbert's original purpose in introducing it) and it is utilized for example in Bourbaki's axiomatic set theory [4] in the guise of τ, where it yields freely the axiom of choice, the ε-symbol has not been considered very much in model theory perhaps because its additional expressive power in the context of first order logic seems null. It could be argued that it is just a clever notation for handling Skolem functions. The ε-symbol allows us to define the first order quantifiers: $\exists x\,\varphi(x) \equiv \varphi(\varepsilon x\,\varphi)$, but no new Lindström quantifier is definable in first order logic with its help, as we will see.

This changes radically if we allow the ε-symbol in languages with generalized quantifiers. Consider, for example, the quantifier

M. Krynicki et al. (eds.), *Quantifiers: Logics, Models and Computation, Volume II*, 63-78.

$$Q_1^E xy\, \varphi(x,y) \;\equiv\; \text{"}\varphi \text{ defines an equivalence relation having uncountably many equivalence classes".}$$

It is well known that it is not definable in $L_{\omega\omega}(Q_1)$, logic with the quantifier "there are uncountably many ...". It is not even definable in $L_{\infty\omega}(Q|Q \in \mathcal{C})$ if $\mathcal{C}$ consists of monadic type quantifiers, ordering quantifiers, or many other natural kinds of quantifiers. However, it is defined in $L_{\omega\omega}(Q_1)$ enriched with ε by the formula:

$$eq(\varphi(x,y)) \wedge Q_1 y(y = \varepsilon x\, \varphi(x,y)),$$

where $eq(\varphi)$ says that $\varphi(x,y)$ defines an equivalence relation, because in such case $Q_1 y(y = \varepsilon x\, \varphi(x,y))$ says that there is an uncountable set of representatives for the relation.

It is the main purpose of this note to show that the situation described above is very general. For a large class of quantifiers the expressive power of the logics they generate increases when the ε-symbol is allowed. Given a regular logic L, let εL be the extension of L obtained by adding to L all Lindström-Mostowski quantifiers definable with the help of ε. We will see that εL is a regular extension of L, preserving compactness, axiomatizability, and other model theoretical properties (§2). A measure of the strength of εL is given by the next inequality, where qL is the congruence closure of L introduced in [12], and ΔL is the closure under Δ-interpolation:

$$qL \leq \varepsilon L \leq \Delta L.$$

Hence, $\varepsilon L \equiv L$ in case L satisfies the interpolation theorem, and $L < \varepsilon L$ if L is a proper extension of $L_{\omega\omega}$ generated by monadic type quantifiers, ordering quantifiers, or any of what we call thin quantifiers in [5], because in such case $L < qL$.

In order to study εL we study first the most general logic L^ε obtained by combining freely the logical operations of L and the formation of ε-terms. Its natural semantics is given by second order structures (A, F), with A a first order structure, and $F : P(\mid A \mid) \to \mid A \mid$ a choice function in the nonempty subsets of the domain $\mid A \mid$ of A which interprets ε. The logic εL consists precisely of those sentences in L^ε whose meaning is independent of the choice function interpreting ε. We show that L^ε inherits most model theoretic properties of L as axiomatizability, compactness, interpolation, and Beth's definability theorem (§1), and give a criterion for the elimination of ε (§3). Finally, we discuss briefly more general term operators (§4).

Some of our results have been shown before for the special case of $L_{\omega\omega}(Q_1)$ by Sánchez in [13]. The model theory of first order logic with term operators more general than ε has been studied by Scott in [14], Corcoran, Hatcher and Herring in [6], and da Costa in [7].

1. The Logic L^ε

Let $L = L_{\omega\omega}(Q|Q \in \mathcal{C})$ where the Q are Lindström-Mostowski quantifiers, or more generally let $L = L_{\mathcal{A}}(Q|Q \in \mathcal{C}) = L_{\infty\omega}(Q|Q \in \mathcal{C}) \cap \mathcal{A}$, where $\mathcal{A}$ is an admissible set in the sense of [3]. We associate to L a logic L^ε in the following way. The language of formulae and terms of $L^\varepsilon(\tau)$ is built inductively as usual from variables, function and relation symbols of τ, first order operations, the quantifiers in $\mathcal{C}$, conjunctions $\wedge\Phi$ whenever $\Phi \in \mathcal{A}$, plus the formation of *ε-terms*:

> If φ is a formula of $L^\varepsilon(\tau)$ then $\varepsilon x\, \varphi$ is a term of $L^\varepsilon(\tau)$ where the variable x is bound.

$\varepsilon x\, \varphi$ is treated syntactically as an ordinary term, it may be substituted in any other term or atomic formula. The semantics of $L^\varepsilon(\tau)$ is given by pairs (A, F) wich we will call *ε-structures*, where A is a first order structure of type τ, $F : P(|A|) \to |A|$ is a choice function in nonempty sets, and $F(\emptyset)$ is arbitrary. The evaluation of terms is extended to ε-terms by declaring inductively:

$$[\varepsilon x\, \varphi]^{(A,F)} = F[\varphi(x)^{(A,F)}]$$

where $\varphi(x)^{(A,F)} = \{a \in A|(A,F) \models \varphi(x/a)\}$. We will use boldface letters to denote sequences of variables or elements, and we will write $\varphi(x, \boldsymbol{b}]^{(A,F)}$ for $\varphi(x, \boldsymbol{y})^{(A,\boldsymbol{b},F)}$, where $\boldsymbol{b}$ interprets $\boldsymbol{y}$.

For appropriate definitions of isomorphism and reduct, L^ε satisfies the basic axioms for logics (see Definition 1.1.1 in [8]). By construction L^ε is closed under substitutions, which yields closure under first order logical operations, and all quantifiers definable in L. The relativization of an ε-structure (A, F) to a subuniverse U closed under the functions of A and containing $F(\emptyset)$ may be defined in the obvious way: $(A,F) \restriction U = (A \restriction U, F \restriction P(U))$. If L is closed under relativizations so is L^ε, because one may define inductively relativizations t^P, φ^P for terms and formulae of L^ε, in particular: $[\varepsilon x\, \varphi]^P = \varepsilon x(\varphi^P)$.

The following schemes are valid in L^ε:

A1. $\forall x(\varphi \leftrightarrow \psi) \to \varepsilon x\, \varphi = \varepsilon x\, \psi$

A2. $\exists x\, \varphi(x, \boldsymbol{y}) \to \varphi(\varepsilon x\, \varphi(x, \boldsymbol{y}), \boldsymbol{y})$

A3. $\varepsilon x\, \varphi(x, \boldsymbol{y}) = \varepsilon z\, \varphi(z, \boldsymbol{y})$, if z is free for x in $\varphi(x, \boldsymbol{y})$, and they form a strongly complete deductive system for $L^{\varepsilon}_{\omega\omega}$ when added to first order inference rules, [A]. We will show that this follows from a more general fact: axiomatizability as well as most other model theoretic properties of an arbitrary logic L are inherited by L^{ε}. For this purpose, and generalizing a method utilized in [6], we interpret L^{ε} in L.

DEFINITION. Extend each type τ to a type τ^* by introducing an n-ary function symbol $f_{\psi(x,\boldsymbol{y})}$ for every formula $\psi = \psi(x, y_1, \ldots, y_n)$ of $L^{\varepsilon}(\tau)$, with all its free variables displayed and the variable x distinguished. When x and ψ are clear from the context we will write simply f_ψ instead of $f_{\psi(x,\boldsymbol{y})}$. Then define transformations:

i) $(A, F) \mapsto (A, F)^* = (A, F_\psi)_{\psi \in L(\tau)}$ from ε-structures of type τ into first order structures of type τ^* by taking: $F_\psi(\boldsymbol{a}) = F[\psi(x, \boldsymbol{a})^{(A,F)}]$.

ii) $\varphi \mapsto \varphi^*$ from formulae of $L^{\varepsilon}(\tau)$ into $L(\tau^*)$, where φ^* is the result of substituting the *most superficial* subterms $\varepsilon x\, \psi(x, \boldsymbol{y})$ of φ (this is those not in the scope of any other ε-term) by the atomic term $f_{\psi(x,\boldsymbol{y})}(\boldsymbol{y})$.

iii) $t \mapsto t^-, \sigma \mapsto \sigma^-$ from terms and formulae of $L(\tau^*)$ into $L^{\varepsilon}(\tau)$, by changing inductively the function symbols f_θ to the corresponding ε-terms. The only nontrivial clause of the inductive definition is:

$$[f_{\theta(x,y^1,\ldots,y^n)}(t_1, \ldots, t_n)]^- = \varepsilon x'\, \theta^{-\prime}(x', t_1^-, \ldots, t_n^-),$$

where $\theta^{-\prime}$ is a variant of θ^- chosen so that all the free variables of the term t_i^- are free for y_i in $\theta^{-\prime}$, and x' does not occur in the t_i^- and is free for x in $\theta^{-\prime}(x, \boldsymbol{y})$. For example, if $\psi(x, \boldsymbol{y})$ is the formula $\forall w(x = \varepsilon z\, R(x, z, y, w))$ and $\theta(x, y, w)$ is the formula $P(x, y, w)$ then the term $f_\psi(f_\theta(z, w))$ becomes

$$\varepsilon x'[\forall w'(x' = \varepsilon z'\, R(x', z', \varepsilon x\, P(x, z, w), w'))].$$

LEMMA 1. i) $(A, F) \models \varphi$ *if and only if* $(A, F)^* \models \varphi^*$.
ii) $(A, F) \models \varphi^-$ *if and only if* $(A, F)^* \models \varphi$.

PROOF. i) Trivial.

ii) One shows simultaneously by induction the statements $\psi^-(x, \ldots)^{(A,F)} = \psi(x, \ldots)^{(A,F)^*}$ and $t^{-(A,F)} = t^{(A,F)^*}$ for any formula ψ and term t in $L(\tau^*)$. The only nontrivial step is the following:

$$\begin{aligned}
[f_\theta(t_1, \ldots, t_n)]^{-(A,F)} &= [\varepsilon x'\, \theta^{-\prime}(x', t_1^-, \ldots, t_n^-)]^{(A,F)} \\
&= F[\theta^{-\prime}(x', t_1^{-(A,F)}, \ldots, t_n^{-(A,F)})^{(A,F)}] \\
&= F[\theta(x, t_1^{(A,F)^*}, \ldots, t_n^{(A,F)^*})^{(A,F)}] \\
&= f_\theta^{(A,F)^*}(t_1^{(A,F)^*}, \ldots, t_n^{(A,F)^*}) = [f_\theta(t_1, \ldots, t_n)]^{(A,F)^*}
\end{aligned}$$

because $\theta^{-\prime}(x', \ldots]^{(A,F)} = \theta^{-}(x, \ldots]^{(A,F)}$, and the induction hypothesis applied to $\theta(x, \ldots)$, and the t_i. □

DEFINITION. For a sentence φ, let $H(\varphi)$ be the conjunction of the universal closures of the formulae

A1*. $\forall x\,(\psi^*(x,\boldsymbol{y}) \leftrightarrow \theta^*(x,\boldsymbol{y})) \rightarrow f_\psi(\boldsymbol{y}) = f_\theta(\boldsymbol{y})$

A2*. $\exists x\,\psi^*(x,\boldsymbol{y}) \rightarrow \psi^*(f_\psi(\boldsymbol{y}),\boldsymbol{y})$

A3*. $f_{\psi(x,\boldsymbol{y})}(\boldsymbol{y}) = f_{\psi(z,w)}(\boldsymbol{y})$, if z is free for x, and w_i is free for y_i in $\psi(x,\boldsymbol{y})$.

where $\psi(x,\boldsymbol{y})$ and $\theta(x,\boldsymbol{y})$ run through the subformulae of φ. For a set of formulae Σ, let $H(\Sigma) = \{H(\varphi) | \varphi \in \Sigma\}$.

Notice that for the logics we are considering $\varphi \in L^\varepsilon$ implies $H(\varphi) \in L$. In the case of infinitary logics this follows from the closure properties of admissible sets, in particular closure under Σ-recursive functions (see [3]).

LEMMA 2. *Let Σ be a class of formulae of L closed under subformulae and let G_θ be a choosing of interpretations in a first order structure A for the f_θ, $\theta \in \Sigma$. Then $(A, G_\theta)_{\theta\in\Sigma} \models H(\Sigma)$ if and only if there is a choice function F in $P(|A|)$ such that $G_\theta = F_\theta$ for all $\theta \in \Sigma$.*

PROOF. Assume $(A, G_\theta)_{\theta\in\Sigma} \models H(\Sigma)$, and define for all $S \in P(|A|)$:

$$F(S) = \begin{cases} G_{\psi(x,\boldsymbol{y})}(\boldsymbol{a}) & \text{if } S = \psi^*(x,\boldsymbol{a}]^{(A,G_\theta)_\theta} \text{ for some} \\ & \psi(x,\boldsymbol{y}) \in \Sigma, \text{ and } n\text{-tuple } \boldsymbol{a} \text{ in } A. \\ \text{arbitrary choice in } S, & \text{otherwise.} \end{cases}$$

F is well defined because the G_ψ satisfy A1*, A3*, and it is a choice function by A2*. Since Σ is closed under subformulae it is possible to prove by induction that for all $\psi(x,\boldsymbol{y}) \in \Sigma$ and $b, \boldsymbol{a}$ in A:

$$(1) \qquad (A, G_\theta)_{\theta\in\Sigma} \models \psi^*[b,\boldsymbol{a}] \Leftrightarrow (A,F) \models \psi[b,\boldsymbol{a}],$$

this implies immediately $F_\psi(\boldsymbol{a}) = F(\psi(x,\boldsymbol{a}]^{(A,F)}) = F(\psi^*(x,\boldsymbol{a}]^{(A,F_\theta)_\theta}) = G_\psi(\boldsymbol{a})$. For atomic ψ, (1) is obvious because then $\psi^* = \psi$ and the formula depends only on A. Assume (1) holds fo all proper subformulae of ψ, then $(A, G_\theta)_{\theta\in\Sigma} \models \psi^*[b,\boldsymbol{a}]$ if and only if $(A, G_\theta)_{\theta\in Subf(\psi)-(\psi)} \models \psi^*[b,\boldsymbol{a}]$ since the symbol f_θ appears in ψ^* only for proper subformulae θ of ψ, if and only if $(A, F_\theta)_{\theta\in\Sigma-(\psi)} \models \psi^*[b,\boldsymbol{a}]$ since by induction hypothesis $F_\theta = G_\theta$ for proper subformulae θ of ψ, if and only if $(A,F) \models \psi[b,\boldsymbol{a}]$ by Lemma 1.

The other direction is trivial since schemes A1*, A2*, and A3* are valid in any $(A,F)^*$, the first two because they are the * translations of A1, A2, and the third because $\psi(x,y)^{(A,F)} = \psi(z,w)^{(A,F)}$. As θ is in Σ for any f_θ in $\varphi \in \Sigma$, then $(A,F_\theta)_{\theta\in\Sigma} \models H(\varphi)$. □

THEOREM 1. i) *If L is compact (respectively (α,β)-compact, $[\alpha,\beta]$-compact) so is L^ε.*

ii) *L^ε has the same Hanf and Löwenheim numbers as L.*

iii) *If $L = L\omega\omega(Q|Q \in \mathcal{C})$ is recursively axiomatizable, so is L^ε.*

iv) *If L is axiomatizable by a system of schemes, then L^ε is axiomatized by the same schemes (generalized to admit the ε-formulae) plus schemes* A1, A2, A3.

PROOF. i) Let Σ be the closure of T under subformulae. If every finite subset of a theory $T \subseteq L^\varepsilon(\tau)$ has a model, the same is true of each finite subset of $\{\varphi^*|\varphi \in T\} \cup H(\Sigma)$ by Lemma 1. By compactness of L, the last theory has a model $(A,G_\theta)_{\theta\in\Sigma}$, which by Lemma 2 induces a model (A,F) of T. The preservation of (α,β) and $[\alpha,\beta]$-compactness is similar, utilising that T and $\{\varphi^*|\varphi \in T\} \cup H(\Sigma)$ are both finite or have the same cardinality.

ii) By the last remark in the proof of (i).

iii) If $L(\tau)$ has a recursively enumerable syntax so do $L^\varepsilon(\tau)$ and $L(\tau^*)$. Given $\varphi \in L(\tau)$, the formula $H(\varphi) \to \varphi^*$ belongs to $L(\tau^*)$ and may be obtained recursively from φ because it has finitely many subformulae. Moreover $\varphi \in L^\varepsilon$ is valid if and only if $H(\varphi) \to \varphi^*$ is valid by Lemmas 1,2. Enumerating recursively the validities of $L(\tau^*)$ of the form $H(\varphi) \to \varphi^*$, we may detect and generate the validities φ of $L^\varepsilon(\tau)$.

iv) If φ is valid, let $\beta_1,\ldots,\beta_n$ be a proof of $H(\varphi) \to \varphi^*$ in the system of schemes of L, then $\beta_1^-,\ldots,\beta_n^-$ is a proof with the same schemes of $H(\varphi)^- \to \varphi^{*-}$. But $H(\varphi)^-$ is a conjunction of cases of schemas A1–A3, and $\varphi^{*-} = \varphi$. □

REMARK. Generalizing part (iii) of the above theorem: if $L = L_{\mathcal{A}}(Q|Q \in \mathcal{C})$ then $Val(L^\varepsilon)$ is Δ-definable in $\mathcal{A}$ over $Val(L)$. Part (iv) holds for infinitary logics axiomatizable by schemes, for example $L\omega_1\omega$.

THEOREM 2. *If L satisfies interpolation, so does L^ε. The same holds for Δ-interpolation, and for Beth's definability theorem in case $L = L\omega\omega(Q|Q \in \mathcal{C})$.*

PROOF. i) Let $K_i = \{(A,F) | \exists R_i(A,R_i,F) \models \varphi_i\}$, $\varphi_i \in L^\varepsilon(\tau_i), i = 1,2$, be disjoint PC classes of ε-structures of type $\tau = \tau_1 \cap \tau_2$, where R_i is a sequence of first order relations. Let $\Sigma = (Subf(\varphi_1) \cup Subf(\varphi_2)) \cap L(\tau)$ and define for $i = 1,2$:

$$K_i^* = \{(A, G_\theta)_{\theta\in\Sigma} | \exists R_i \exists G_{\mu,\mu\in Subf(\varphi_i)-\Sigma}(A, R_i, G_\theta, G_\mu) \models H(\varphi_i) \wedge \varphi_i^*\},$$

then the K_i^* are disjoint PC classes of L, because if we had $(A, G_\theta)_{\theta\in\Sigma} \in K_i^* \cap K_2^*$ then for some $R_i, G_{\mu_i}, \mu_i \in Subf(\varphi_i) - \Sigma, i = 1,2$:

$$(A, R_i, G_\theta, G_{\mu_i})_{\theta\in\Sigma,\mu_i\in Subf(\varphi_i)-\Sigma} \models H(\varphi_i) \wedge \varphi_i^*$$

and so

$$(A, R_1, R_2, G_\rho)_{\rho\in Subf(\varphi_1)\cup Subf(\varphi_2)} \models (H(\varphi_1) \wedge H(\varphi_2)) \wedge \varphi_1^* \wedge \varphi_2^*.$$

By Lemmas 1 and 2 there is F such that $F_\rho = G_\rho, \rho \in Subf(\varphi_1) \cup Subf(\varphi_2)$, and $(A, R_1, R_2, F) \models \varphi_1 \wedge \varphi_2$; hence, $(A,F) \in K_1 \cap K_2$, which is absurd. Now, let σ be an interpolant of K_1^* and K_2^* in $L(\tau \cup \{f_\theta : \theta \in \Sigma\})$, then σ^- is an interpolant of K_1 and K_2 in $L^\varepsilon(\tau)$ by Lemmas 1,2.

ii) To see that Δ-interpolation is preserved, notice that if K_1 and K_2 are complementary PC classes of $L^\varepsilon(\tau)$ then K_1^* and $K_2^* \cup \{(A, G_\theta)_{\theta\in\Sigma} | (A, G_\theta)_{\theta\in\Sigma} \models \neg \wedge H(\Sigma)\}$ are also complementary PC of type τ^*. The rest of the proof is the same.

iii) First notice that the ordinary Beth's theorem implies the following simultaneous version: if $\psi(R_1, \ldots, R_n) \in L(\tau \cup \{R_1, \ldots, R_n\})$ is such that

$$\psi(R_1, \ldots, R_n) \wedge \psi(R_1', \ldots, R_n') \models \forall x_i(R_i(x_i) \leftrightarrow R_i'(x_i)), i = 1, \ldots, n,$$

then there are $\sigma_i \in L(\tau)$ such that $\psi(R_1, \ldots, R_n) \models \forall x_i(R_i(x_i) \leftrightarrow \sigma_i(x_i))$. Suppose now that $\varphi(R) \in L^\varepsilon(\tau \cup \{R\})$ defines implicitly R, then

$$(2) \qquad H(\varphi(R)) \wedge H(\varphi(R')) \wedge \varphi(R)^* \wedge \varphi(R')^* \models \forall x(R(x) \leftrightarrow R'(x)), \forall y(f_\theta(y) = f_{\theta'}(y))$$

for each $\theta \in Subf(\varphi)$, with $\theta' = \theta(R/R')$. To see this assume

$$\mathcal{B} = (A, R, R', G_\mu, G_\theta, G_{\theta'})_{\mu\in Subf(\varphi)\cap L(\tau),\theta\in Subf(\varphi)-L(\tau)}$$

satisfies the left hand formula in (2), then by Lemmas 1, 2 there is F such that $F_\rho = G_\rho$ for $\rho \in Subf(\varphi) \cup Subf(\varphi')$ and $\mathcal{B}^+ = (A, R, R', F) \models \varphi(R) \wedge \varphi(R')$. By hypothesis $\mathcal{B}^+ \models \forall x(R(x) \leftrightarrow R'(x))$ and so $G_\theta = F_\theta = F_{\theta'} = G_{\theta'}$; hence $\mathcal{B} \models \forall y(f_\theta(y) = f_{\theta'}(y))$. Using that φ has finitely many subformulae and applying the simultaneous version of Beth's theorem in L, we have $\sigma \in L(\tau^*)$ such that $\models H(\varphi(R)) \wedge \varphi(R)^* \to \forall x(R(x) \leftrightarrow \sigma(x))$, and so by Lemma 1, $\models \varphi(R) \to \forall x(R(x) \leftrightarrow \sigma^-(x))$. □

EXAMPLE. i) $L\omega_1\omega^\varepsilon, L\omega_1\omega^\varepsilon$, and any countable admissible fragment of $L\omega_1\omega^\varepsilon$ satisfy the interpolation theorem. Moreover, they are axiomatizable for validities by their usual shemes and rules, plus A1, A2, A3.

ii) $L\omega\omega(Q_1)^\varepsilon$ is countably compact and it is axiomatized by Keisler's axioms for $L\omega\omega(Q_1)$ plus A1, A2, A3. This fact has been shown before in [13].

2. The ε-Closure of a Logic, ε-Invariant Formulae

L^ε is not a logic for first order structures in the sense of Lindström since the satisfiability of an ε-sentence in a first order structure may depend not only on the fact that ε represents a choice function "in general" but also whether the choice function satisfies or not some additional mathematical conditions. This is the case of the following pure equality sentence of $L\omega\omega^\varepsilon$:

$$\forall x \forall y(x \neq y \to (\varepsilon z(z \neq y) \neq \varepsilon z(z \neq x)),$$

for which we can not determine the truth value in a set A until we specify the choice function interpreting ε. This fact could be construed as implying that ε is not a logical but a mathematical concept from the perspective of first order structures.

However, in the context of some particular formulae ε behaves as a genuine logical concept, representing "some choice function no matter which", a *generic choice function*. For example, the sentence

$$\forall x(\varepsilon y(y \neq x) = x)$$

holds in a structure (A, F) if and only if $|A| = 1$, independently of which is F. Similarly, if φ does not contains ε, the truth value of any sentence of the form $\varphi(\varepsilon x \varphi)$, or the sentence $eq(\varphi(x,y)) \wedge Q_1 y(y = \varepsilon x\, \varphi(x,y))$ discussed in the introduction, may be found in a first order structure by interpreting ε by any choice function we like. This ambiguous but invariant meaning of ε seems to be the original conception of Hilbert and Ackermann, and justifies the next definition.

DEFINITION. A formula $\varphi \in L^\varepsilon$ will be called *ε-invariant* if and only if for any structure A of type τ and choice functions F, G in $P(|A|)$ we have

$$(A, F) \models \varphi \text{ if and only if } (A, G) \models \varphi.$$

$\varepsilon L(\tau)$ will be the subclass of $L^\varepsilon(\tau)$ consisting of the ε-invariant formulae. We can make εL a Lindström logic by defining for a first order structure A of type τ, and $\varphi \in \varepsilon L(\tau) : A \models_{\varepsilon L} \varphi$ if and only if $\exists F(A, F) \models \varphi$.

It is straightforward to check that the basic Lindström axioms (see Definition 1.1.1 in [8]) hold for εL. Evidently, $L \leq \varepsilon L$. The definability of the syntax of εL is problematic because it depends on validity; εL is not even closed under subformulae. However, εL inherits the main model theoretic properties of L.

THEOREM 3. i) *If L is a regular logic so is* εL.
ii) *If L is recursively axiomatizable, so is* εL.
iii) *If L is compact (respectively* (α,β)*-compact,* $[\alpha,\beta]$*-compact) so is* εL.

PROOF. i) Since L^ε is regular when L is, it is enough to check that εL is closed in L^ε under substitutions and relativizations. If $\theta(R_1, \ldots, R_n)$ and $\psi_1, \ldots, \psi_n$ are invariant formulae of L^ε then the substituted formula $\theta(\psi_1, \ldots, \psi_n)$ belongs to L^ε. Now, given A, F, G we have by hypothesis $\psi_i(\boldsymbol{x}_i)^{(A,F)} = \psi_i(\boldsymbol{x}_i)^{(A,G)}, i = 1, \ldots, n$, and so by definition and invariance of θ: $(A, F) \models \theta(\psi_1, \ldots, \psi_n) \Leftrightarrow (A, \psi_1^{(A,F)}, \ldots, \psi_n^{(A,F)}, F) \models \theta(R_1, \ldots, R_n) \Leftrightarrow (A, \psi_1^{(A,G)}, \ldots, \psi_n^{(A,G)}, G) \models \theta(R_1, \ldots, R_n) \Leftrightarrow (A, G) \models \theta(\psi_1, \ldots, \psi_n)$; hence, $\theta(\psi_1, \ldots, \psi_n)$ is ε-invariant. Similarly, if φ is ε-invariant and P is a unary predicate symbol then φ^P is ε-invariant because $(A, F) \models \varphi^P \Leftrightarrow (A \restriction P^A, F \restriction P^A) \models \varphi \Leftrightarrow (A \restriction P^A, G \restriction P^A) \models \varphi \Leftrightarrow (A, G) \models \varphi^P$.

ii) If $L(\tau)$ is recursively enumerable for validities then $\varepsilon L(\tau)$ has a recursively enumerable syntax, because φ is ε-invariant if and ony if

$$\models H(\varphi) \wedge H(\varphi)^0 \to (\varphi^* \leftrightarrow \varphi^{*0}),$$

where ψ^0 denotes a renaming of ψ with respect to the function symbols $f_\theta, \theta \in Subf(\varphi)$. Now, by Theorem 1, $Val(L^\varepsilon)$ is recursively enumerable, hence $Val(\varepsilon L) = \varepsilon L(\tau) \cap Val(L^\varepsilon)$ is recursively enumerable.

iii) Follows from the fact that εL is a sublogic of L^ε with respect to ε-structures, and Theorem 1. □

Recall that a logic L is *congruence closed* [12] if for every $\varphi \in L(\tau)$, $\tau = \langle R_1, \ldots, R_n \rangle$, it is closed under the quantifier $Q^{\varphi,\tau}$ of type τ where:

$$A \models Q^{\varphi,\tau} \boldsymbol{x}_1 \ldots \boldsymbol{x}_n yz(\varphi_1(\boldsymbol{x}_1), \ldots, \varphi_n(\boldsymbol{x}_n), \theta(y, z))$$

if and only if θ^A is a congruence relation in $(|A|, \varphi_1(\boldsymbol{x}_1)^A, \ldots, \varphi_n(\boldsymbol{x}_n)^A)$ and $(|A|, \varphi_1(\boldsymbol{x}_1)^A, \ldots, \varphi_n(\boldsymbol{x}_n)^A)/\theta^A \models \varphi$.

For any logic L there is a smallest congruence closed extension qL of L, called the *congruence closure* of L.

THEOREM 4. *$\varepsilon L \leq \Delta L$ for any logic L. If L has relativizations then εL is congruence closed and so $qL \leq \varepsilon L$.*

PROOF. If φ is invariant so is $\neg\varphi$; hence, by Lemma 2

$$A \models \varphi \text{ if and only if } \exists G_{\theta,\theta \in Sub(\varphi)}(A, G_\theta, \ldots) \models H(\varphi) \wedge \varphi^*$$

$$A \models \neg\varphi \text{ if and only if } \exists H_{\theta,\theta \in Sub(\varphi)}(A, H_\theta, \ldots) \models H(\neg\varphi) \wedge \neg\varphi^*,$$

this shows $\varepsilon L \leq \Delta L$. Now, if L has relativizations, εL has relativizations also, and it is anyway closed under substitutions, then $Q^{\varphi,\tau}$ is definable in εL by:

$$Q^{\varphi,\tau}\boldsymbol{x}_1 \ldots \boldsymbol{x}_n xy(\psi_1(\boldsymbol{x}_1), \ldots, \psi_n(\boldsymbol{x}_n), \theta(x,y)) \equiv$$

$$Congr(\theta(x,y), \psi_1(\boldsymbol{x}_1), \ldots, \psi_n(\boldsymbol{x}_n)) \wedge \varphi(\psi_1, \ldots, \psi_n)^{\{x | x = \varepsilon y \theta(x,y)\}}$$

where $Congr(\theta(x,y), \psi_1(\boldsymbol{x}_1), \ldots, \psi_n(\boldsymbol{x}_n))$ is the first order formula saying that the relation defined by θ is a congruence with respect to the relations defined by the ψ_i. The equivalence holds because the quotient of a structure by a congruence relation θ is isomorphic to the relativization of said structure to any set of representatives of the relation, and $\{x | x = \varepsilon y\, \theta(x,y)\}$ defines such a set of representatives since $\theta(a,b)$ implies $\forall y(\theta(a,y) \leftrightarrow \theta(b,y))$ and so $\varepsilon y\, \varphi(a,y) = \varepsilon y\, \varphi(b,y)$ by A1. □

COROLLARY 1. *Let L be a logic satisfying Δ-interpolation, then $\varepsilon L \equiv L$.*

EXAMPLE. No new Lindström-Mostowski quantifiers are definable with the help of ε in $L\omega\omega$, $L\omega_1\omega$, or any admissible fragment of $L\omega_1\omega$.

The fact the εL contains the congruence closure has the consequence that it extends properly L for a very large class of logics. We say that a quantifier Q is *thin* if there is κ such that for each $\delta \geq \kappa$ it is definable in structures of power δ by a sentence of finite quantifier rank of $L_{\infty\omega}(\mathbf{T}_\delta)$, where $(A, R_1, \ldots, R_n) \in Q \in \mathbf{T}_\delta$ implies A does not contain subsets of power δ homogenous for some R_i. This is a complicated definition but the important point is that all monadic type quantifiers, all ordering quantifiers, and many other natural quantifiers are thin (see [5]).

COROLLARY 2. *If* $L = L\omega\omega(Q|Q \in \mathcal{C})$ *is a proper extension of* $L\omega\omega$ *closed under relativizations and* $\mathcal{C}$ *consists of thin quantifiers then* $L < \varepsilon L$.

PROOF. By Theorem 1 in [5] we must have $L < qL$. Then use Theorem 4. □

EXAMPLE. $L(Q_\alpha) < \varepsilon L(Q_\alpha)$ and $L(Q_\kappa^{cof}) < \varepsilon L(Q_\kappa^{cof})$ for any ordinal α, and regular cardinal κ. Notice that $\varepsilon L(Q_0) \leq \Delta L(Q_0) = L\omega_1^{CK}\omega$ and $\varepsilon L(Q_\omega^{cof}) \leq \Delta L(Q_\omega^{cof}) \leq L\omega\omega(aa)$.

Interpolation and Δ-interpolation are trivially inherited by εL from L by Corollary 1. This is not immediate for Beth's theorem, which does not imply Δ-interpolation.

THEOREM 5. $L = L\omega\omega(Q|Q \in \mathcal{C})$ *satisfies Beth's definability theorem, then so does* εL.

PROOF. If $\varphi(R) \in L^\varepsilon$ defines implicitly R and is ε-invariant, find $\psi(x) \in L^\varepsilon$ defining explicitly R (it exists by Theorem 2). Then it is easy to see that $\psi(x) \wedge \psi(R/\psi(x))$ is ε-invariant and defines explicitly R. □

3. Elimination of ε

Given a theory T in L^ε, one could ask for which $\varphi \in L^\varepsilon$ there is a formula $\varphi' \in L$ such that $T \models \varphi \leftrightarrow \varphi'$. This problem has been considered by Scott [14] in the case of first order logic with a term operator more general than ε (see next section). If L satisfies Δ-interpolation and T is the empty theory, an answer follows immediately from Corollary 1: a sentence $\varphi \in L^\varepsilon$ is equivalent to a sentence of L if and only if it is ε-invariant. For the general case we need stronger conditions on the logic.

A sentence $\varphi \in L^\varepsilon$ will be called *T-invariant* if for any $(A, F), (A, G) \models T : (A, F) \models \varphi$ iff $(A, G) \models \varphi$. So $\emptyset$-invariance is the same as ε-invariance.

THEOREM 6. *Let* L *be a logic satisfying interpolation and compactness, let* $T \subset L^\varepsilon$ *and* $\varphi \in L^\varepsilon$. *Then there is* $\varphi' \in L$ *such that* $T \models \varphi \leftrightarrow \varphi'$ *if and only if* φ *is T-invariant.*

PROOF. The condition is obviously necessary. Now, if φ is T-invariant, the classes

$$K_i = \{A | \exists F_{\theta, \theta \in Sub(T \cup \{\varphi\})} (A, F_\theta)_\theta \models H(T \cup \{\varphi\}) \wedge T^* \wedge \psi_i\},\ i = 1, 2$$

with $\psi_1 = \varphi^*$ and $\varphi_2 = \neg\varphi^*$ are disjoint PC_Δ in L and by compactness they are contained in disjoint PC classes. If δ is a interpolant of these two classes, say $K_i \subseteq Mod(\delta)$, then we have $T \models \varphi \leftrightarrow \delta^-$ by Lemmas 1,2. □

REMARK. If T is finite compactness is unnecessary in the above theorem. If T has the property that for any first order structure A there is F with $(A, F) \models T$, (for example when $T \subset L$), then Δ-interpolation instead of full interpolation is enough for the above theorem to hold, because the PC classes may be extended to complementary PC classes.

The only familiar logic satisfying compactness and interpolation is $L\omega\omega$. Applying Theorem 6 to $L\omega\omega$ we obtain a result due to Scott.

COROLLARY 3. (Scott [14]) *Let T be a theory of $L\omega\omega^\varepsilon$ such that for any models $(A,F),(B,G)$ of $T, A \approx B$ implies $(A,F) \equiv_{L^\varepsilon} (B,G)$, then for any $\varphi \in L^\varepsilon$ there is $\varphi' \in L$ such that $T \models \varphi \leftrightarrow \varphi'$.*

PROOF. The condition given is necessary for elimination of ε, and it implies that any φ is T-invariant. □

4. Other Term Operators

Scott [14] and later Corcoran, Hatcher, and Herring [6] discussed the logic $L\omega\omega^v$ where v is general term operator similar to ε but satisfying only A1 and A3:

A1(v). $\forall x(\varphi \leftrightarrow \psi) \rightarrow vx\,\varphi = vx\,\psi$

A2(v). $vx\,\varphi(x,y) = vz\,\varphi(z,y)$, if z is free for x in $\varphi(x,y)$.

They show the completeness of these two schemes with respect to *v-structures* (A,F) where F is now any function $F : P(|A|) \rightarrow |A|$. Scott works with a two sorted version of v-structures in which the value $vx\,\varphi(x)$ may be out of A, in a universe of virtual individuals, and shows in this context Corollary 3 above. Our general results for L^ε hold with almost identical proofs for the logic L^v which results by adding v to L.

THEOREM 7. *L^v inherits from L compactness, Hanf and Löwenheim numbers axiomatizability, and definability properties. Moreover, L^v satisfies the v-elimination analogues of Theorem 6 and Corollary 3.*

However, from our point of view the operator v is not very interesting since it does not have any capability to define new Lindström-Mostowski quantifiers. Let vL be the logic of v-invariant sentences of L^v, this is, invariant with respect to all v-structures.

THEOREM 8. *$vL \equiv L$ for any logic L.*

PROOF. Let $\varphi \in L^v$ be v-invariant and let $\varphi^+(x) \in L$ be the result of changing the most superficial occurrences of v-terms in φ by a variable x not in φ. Given $a \in A$, let $F_a : P(|A|) \to |A|$ be the constant function $F(S) = a$. Then $A \models \varphi \Leftrightarrow (A, F) \models \varphi$ for some (all) $F \Leftrightarrow (A, F_a) \models \varphi$ for all $a \in A \Leftrightarrow A \models \forall x \varphi^+(x)$. □

One should consider then term operators satisfying some other schemes, not necessarily A2, in addition to A1, A3. Such is partially done in the context of first order logic in [14] and [7]. The following is a precise definition of general term operators.

A *term operator* v_C is an assignment $C(A) \subset A^{P(A)}$ to each nonempty set A, satisfying the following conditions:

C1. $\forall A\, C(A) \neq \emptyset$

C2. If $g : A \to B$ is a bijection, then $F \in C(A)$ implies $g \circ F \circ g^{-1} \in C(B)$.

It is easy to verify that with this definition any logic L gives rise to a logic L^{v_C}, enriched with new terms $v_C x\, \varphi(x)$. The *v_C-structures* are the pairs (A, F) with A first order and $F \in C(|A|)$, and the terms $v_C x\, \varphi(x)$ are interpreted in (A, F) by $F(\varphi(x)^{(A,F)})$. There is also a corresponding Lindström logic $v_C L$ of v_C-invariant formulae which extends L. The full study of these operators is not less complex than the study of all Lindström-Mostowski quantifiers. None of the results we proved for ε hold in general for the new operators; however, they seem to provide a good alternative to quantifiers in defining logics. We finish with a concrete example.

Consider the term operator defined by $C(A) = \{F \in A^{P(A)} | F(S) \in S$ if and only if $card(S) \geq \omega_1$ or $S = A\}$. This is,

$$\varepsilon_1 x\, \varphi(x) = \begin{cases} \text{some } x \text{ such that } \varphi(x) \text{ if } Q_1 x\, \varphi(x) \text{ or } \forall x \varphi(x) \\ \text{some } x \text{ such that } \neg\varphi(x) \text{ if } \neg Q_1 \varphi(x) \text{ and } \exists x\, \neg\varphi(x) \end{cases}$$

Then Q_1 is definable from ε_1 and first order operations:

$$Q_1 x \varphi(x) \equiv \exists y [\varphi(\varepsilon_1 x(\varphi(x) \wedge x \neq y)) \wedge \varepsilon_1 x\, (\varphi(x) \wedge x \neq y) \neq y]$$

because the right hand side formula means in a ε_1-structure (A, F) that there is $y \in A$ such that $F(\varphi^A - \{y\}) \in \varphi^A - \{y\}$, and this only happens if $\varphi^A - \{y\}$, and so φ^A, is uncountable. Hence,

$$L\omega\omega^{\varepsilon_1} \equiv L\omega\omega(Q_1)^{\varepsilon_1}.$$

Since the defining sentence must be ε_1-invariant then:

$$\varepsilon_1 L_{\omega\omega} \equiv \varepsilon_1 L\omega\omega(Q_1).$$

In fact,

$$(3) \qquad L\omega\omega(Q_1) < \varepsilon_1 L\omega\omega \equiv \varepsilon_1 L\omega\omega(Q_1) \leq \varepsilon L\omega\omega(Q_1) \leq \Delta L\omega\omega(Q_1).$$

The first inequality is proper because the quantifier $Q_1^E xy\,\varphi(x,y)$ is still definable in $\varepsilon_1 L\omega\omega(Q_1)$ by the formula

$$eq(\varphi(x,y)) \wedge [Q_1 x(x = \varepsilon_1 y\,\varphi(x,y)) \vee Q_1 x \neg Q_1 y\,\varphi(x,y)].$$

It may be seen that $\varepsilon_1 L\omega\omega \leq \varepsilon L(Q_1)$ because for an ε_1-invariant formulae φ, $A \models \varphi$ if and only if there is a ε_1-choice F such that $(A, F) \models \varphi$, if and only if there is a ε-choice F' such that $(A, F') \models \varphi^0$, where φ^0 results from performing inductively in φ the substitution of subterms $\varepsilon_1 x\,\theta(x)$ by

$$\varepsilon x\{[\theta(x) \wedge (Q_1 x\,\theta(x) \vee \forall x\,\theta(x))] \vee [\neg\varphi(x) \wedge \neg(Q_1 x\,\theta(x) \vee \forall x\,\theta(x))]\}.$$

Obviously φ^0 must be ε-invariant by ε_1-invariance of φ. We do not know if the last two inequalities in (3) are proper. Finally, notice that ε_1 satisfies the scheme:

$$\text{A4}(\varepsilon_1) \qquad (Q_1 x\,\varphi(x) \vee \forall x\,\varphi(x)) \leftrightarrow \varphi(\varepsilon_1 x\,\varphi(x))$$

and it is not difficult to show that an axiomatization for $L\omega\omega(Q_1)$ when added to A1(ε_1), A3(ε_1), A4(ε_1) provides a complete axiomatization for $L\omega\omega(Q_1)^{\varepsilon_1}$.

5. Questions

1. Given $L < \Delta L$, when do we have $qL = \varepsilon L$? $\varepsilon L = \Delta L$? The last question may be interesting since in all the known nontrivial cases ΔL is not finitely generated by Lindström quantifiers.
2. Is $\varepsilon L\omega\omega$ stronger than $L\omega\omega$ in finite models? Notice that $L\omega\omega < \Delta L\omega\omega$ in finite models, for example the quantifier "even" is definable in $\Delta L\omega\omega$ but not in $L\omega\omega$.
3. Is every monadic type Lindström quantifier definable from some term operator added to first order logic? We have seen that Q_1 is definable in

$\varepsilon_1 L_{\omega\omega}$, and proceeding similarly it is clear that this works for any Q of type $\langle 1 \rangle$ having the property: $(A, S) \in Q$ iff $(A, S \pm \{a\}) \in Q$; for example any cardinality quantifier. It is also true if either $Qx(x = x)$ or $\neg Qx(x = x)$ is valid, for Chang's quantifier for example.

Added in Proof

Professor Jörg Flum has sent to me an example showing that the answer to question 2 above is positive. The precise strength of $\varepsilon L_{\omega\omega}$ in finite models becomes then an interesting problem related to complexity theory.

University of Colombia, Bogota, Colombia

References

[1] G. ASSER, *Theorie der logischen Auswahlfunktionen*, **Zeitschrift für Mathematische Logik und Grundlagen der Mathematik** 3 (1957), pp. 30–68.

[2] W. ACKERMANN, *Megentheoretische Begrundung der Logik*, **Mathematische Annalen** 115 (1937–8), pp. 1–22.

[3] J. BARWISE, **Admissible sets and structures**, Springer Verlag, 1975.

[4] N. BOURBAKI, **Elements des Mathématiques. Libre I (Theorie des ensembles)**, Cap I, II. Hermann, Paris, 1954.

[5] X. CAICEDO, *Definability properties and the congruence closure*, **Archive for Mathematical Logic** 30 (1990), pp. 231–240.

[6] J. CORCORAN, W. S. HATCHER and J. HERRING, *Variable binding term operators*, **Zeitschrift für Mathematische Logic und Grundlagen der Mathematik** 18 (1972), pp. 177–186.

[7] N. D. A. DA COSTA, *Model theoretical approach to variable binfing term operators*, In **Mathematical Logic in Latin America**, North–Holland, 1980, pp. 133–162.

[8] H. D. EBBINGHAUS, *Extended logics: the general framework*, In **Model Theoretic Logics**, J. Barwise, S. Feferman (eds.), Springer Verlag, 1985.

[9] D. HILBERT, **Die Grundlagen der Mathematik**, Address at the Hamburg Mathematical Seminar (translation in **From Frege to Gödel**, J. van Heijenort (ed.), Harward University Press, 1927.

[10] D. HILBERT and P. BERNAYS, **Grundlagen der Mathematik**, Vol. 2, Springer Verlag, 1939.

[11] A. C. LEISENRING, **Mathematical Logic and Hilbert's ε-symbol**, Macdonald Technical and Scientific, London, 1969.

[12] A. H. MEKLER and S. SHELAH, *Stationary logic and its friends. I*, **Notre Dame Journal of Formal Logic** 26 (1985), pp. 129–138.

[13] C. H. SÁNCHEZ, *Aspectos da eliminabilidade dos operadores nominais*, Tese de Doutorado, Universidade Estadual de Campinas, Sao Paulo, 1988.

[14] D. Scott, *Existence and description in formal logic*, In **Bertrand Russell, Philosopher of the Century**, Little, Brown and Co., Boston, 1967.