HOW MUCH LOGIC IS BUILT INTO NATURAL LANGUAGE?

ED KEENAN

Linguistics
UCLA
keenan@humnet.ucla.edu

Query Does knowing a natural language (English, Japanese, Swahili,...) imply knowing any logic?

The Query is reasonable (First Order) Predicate Logic ($PL_=$) is a "Universal Grammar" for the languages of Elementary Arithmetic, Euclidean Geometry, Set Theory, ... It defines their expressions, their semantic interpretations, and proofs, that syntactically characterize the boolean semantic entailment relation.

1. Properties of PL overtly present in Natural Language (NL)

1.1. Function Symbols (F1s, F2s,...) and Naming Expressions (F0s)

PL: $+$ and $\times$ are F2s, squaring ^{2} is an F1: 2, 3, 3^2, (3^2+2), $(3^2+2)^2, \ldots$

NL: kin terms are F1's: *the dean, the mother of the dean, the mother of the mother of the dean,....* These are easier to understand if we vary the function expression: *the wife of the employer of the mother of the dean, etc., ...*

Recursion = the values of a function lie in its domain, so its application iterates Not limited to possessive constructions. In children's rhymes and songs:

Relative clauses *This is the house that Jack built, This is the malt that lay in the house that Jack built, This is the rat that ate the malt that lay in the house ...*

Prepositional phrases *There's a hole in the bottom of the sea, There's a log in the hole in the bottom of the sea, There's a bump on the log in the hole...*

Compositionality meaning of a derived expression a function of those it is derived from: '$(2+3)$' denotes the value of the function denoted by '$+$' at the numbers denoted by '2','3'.

A Fundamental Similarity PL and NL are recursive, compositional systems. They build infinitely many non-synonymous expressions from a finite list.

Leading Question of Md Linguistics: Account for how we produce and understand arbitrarily many novel expressions in NL. Recursion + Compositionality a partial answer

Recursion (self application) is a "statistical accident." Most functions don't iterate: *The height of the dean, #the height of the height of the dean, ...*

1.2. Predicate-Argument Formulas (FMs) /Sentences (Ss)

PL Simple FMs = Predicate + Names. '$2 > 1$', $2^2 = (3+1)$.

NL abundant: P1s $\approx$ *sleeps*; P2s $\approx$ *praises*; P3s $\approx$ *gives*;

Arguments are often **asymmetrically** related: In PL $2 > 1$ and $1 > 2$ both make sense (but differ in truth value). *I wrote that poem* is natural, *That poem wrote me* is nonsense. The first argument of *write* is its **Agent**, the second its **Patient**.

The second argument of a P2 may be referentially bound to the first, but not conversely:

Ben washed/punished himself *Himself (Heself) washed/punished Ben

P2s in NL may fail to be isomorphic. *Ben washed the car* passivizes to *The car was washed by Ben*. But *Ben has a car* does not passivize: **The car is had by Ben*.

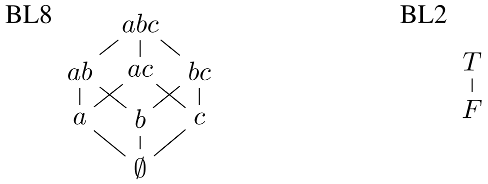
1.3. Boolean operations

In PL *and*, *or*, and *not* build FMs from FMs; they denote boolean functions: *and* is a binary greatest lower bound (glb) operator, noted $x \wedge y$; *or* a binary least upper bound (lub) operator, noted $x \vee y$, *not* is a complement operator, noted $\neg x$. Writing $TV(\phi)$ for the truth value of ϕ, we have $TV(\phi \& \psi) = TV(\phi) \wedge TV(\psi)$, $TV(\phi \,\text{or}\, \psi) = TV(\phi) \vee TV(\psi)$ and $TV(\text{not}\, \phi) = \neg(TV(\phi))$.

Negation Present in all NLs (Dryer 2005). *Bill* isn't *a linguist.*

'and', 'or' and 'neither ... nor ...' combine with expressions in most categories (in PL they only connect FMs); *Both John and Bill either laguhed or cried.*

Boolean Lattices (B,$\leq$) are distributive, complemented lattices. $\leq$ is a boolean partial order:



Here $x \leq y$ iff $x = y$ or you can move up along edges from x to y. $x \wedge y$ is the "highest" element which is $\leq x$ and $\leq y$. $x \vee y$ is the lowest element that both x and y are $\leq$ to. BL2 is the boolean lattice of truth values, which FMs denote.

Gen 1 *The set in which expressions of a category C denote is a boolean lattice (B, $\leq$), supporting that the boolean operations are "properties of mind" (Boole 1854).*

Gen 2 *Modifiers are usually restricting:* tall student $\leq$ student*, that is, all tall students are students, all skillful doctors are doctors, etc.*

Variable usage in NL: *and* may = *and then*, as in *Flo got married and got pregnant* $\neq$ *Flo got pregnant and got married*, or *and as a result*, as in *John drank too much and got sick*. But not always: *Flo is 6 feet tall and studies biology* = *Flo studies biology and is 6 feet tall*. Usage in logic abstracts from this variation to yield $(P\&Q)$

is true iff P is and Q is, whence the semantic symmetry: $P\&Q = Q\&P$. Similarly with *or*, which is sometimes intended as exclusive, as in *John either laughed or cried* (? but not both). But not always: *Do you have two nickels or a dime?* must be answered 'Yes' if you have both two nickels and a dime.

Quantification in PL $Qx\phi$ is a FM, where ϕ a FM, x a variable and Q is either the universal quantifier *all*, noted $\forall$, or the existential quantifier *there exists*, noted $\exists$.

$\forall x(x^2 > x)$ 'The square of any number is greater than that number'

$\exists x(Even(x)\& Prime(x))$ 'There is a number x which is both even and prime'

Semantically $\forall$ an arbitrary glb operator, and $\exists$ an arbitrary lub operator. E.g. $TV(\forall x(x^2 \geq x)) = TV(0^2 \geq 0) \wedge TV(1^2 \geq 1) \wedge (2^2 \geq 2) \wedge \ldots$ Writing $TV(\phi[^x/_b])$ for the truth value of ϕ when the variable x is set to denote b, we see that $\forall$ = "AND writ big"; $\exists$ = "OR writ big." $TV(\forall x\phi) = \bigwedge\{TV(\phi[^x/_b]) \mid b \in E\}$ and $TV(\exists x\phi) = \bigvee\{TV(\phi[^x/_b]) \mid b \in E\}$

PL ties variable binding to quantification. It is enlightening to separate them, as in $ALL(x.\phi)$ where $(x.\phi)$ is a P1 built from a FM (P0) by prefixing the variable x. Then Qs combine directly with P1s to form FMs (P0s). (Read $(x.\phi)$ as $\lambda x.\phi$).

NL Universal Quantification Present in all NLs (knowledgeable conjecture, kc; Stassen 2005, Gil 2005). *All* cats are black; The students have *all* left;

Existential Quantification All NLs may assert and deny existence (kc): *There are / (aren't any) children in the park.*

2. Logical Properties covertly present in NL (Whorf-Sapir Hypothesis)

Knowing English implies knowing the distribution of NPI's (negative polarity items)—e.g., *ever* and *any*, whose presence is licensed by overt negation, as in (1), but also by certain NPs in subject position, as in (2):

(1) a. John hasn't ever been to Pinsk a'. John didn't see any birds on the walk
b.*John has ever been to Pinsk b'.*John saw any birds on the walk.

(2) a. No student here has ever been to Pinsk
a' Neither John nor Mary knew any Russian
b. *Some student here has ever been to Pinsk
b' *Either John or Mary knew any Russian.
c. Fewer than five / *More than five students here have ever been to Pinsk
d. At most / *At least two students here have ever been to Pinsk

query Which NPs license NPI's? What do they have in common with negation?

Gen 3 *NPI licensers are expressions which denote monotone* ***decreasing*** *functions*

Def Let (A,$\leq$) and (B,$\leq$) be posets, F a function from A into B. Then

a. F is *increasing* iff for all $x, y \in$ A, if $x \leq y$ then $F(x) \leq F(y)$.

b. F is *decreasing* iff for all $x, y \in$ A, if $x \leq y$ then $F(y) \leq F(x)$.

Test for Increasingness (↑): if all Ps are Qs and X is a P, therefore X is a Q. Ex: 'some poet' is ↑: Suppose all Londoners drink stout and some poet is a Londoner. Therefore some poet drinks stout.

Gen 4 *Virtually all syntactically underived NPs are ↑: Proper Names (*Ned, Gail*), pronouns (*he, she, they*), demonstratives (*this, those*).*

Gen 5 *The closure of Proper Name denotations under the (complete) boolean operations is the denotation set for all quantified NPs (*No/Most/All students,...*).*

Test for Decreasingness (↓): All Ps are Qs and X is a Q, therefore X is a P.
'No poet' is ↓: if all Londoners drink stout but no poet drinks stout then no poet is a Londoner
Negation is ↓: if Londoner → drinking stout then not drinking stout → not being a Londoner

Gen 6 *The major ways of building NPs from NPs preserve or reverse monotonicity:*

a. *Conjunctions and disjunctions of ↑ NPs are ↑; analogously for ↓ NPs.*
b. *Possessive NPs have the monotonicity value of the possessor: X's doctor is ↑(↓) if X is.*
c. *Negation reverses monotonicity: not more than two boys is ↓ since more than two boys is ↑*

query Which NPs occur naturally in partitives, as in *Two of* _?
yes: Two of *those cats*, two of *John's/the ten/John's ten/my cats*
no: *two of *most cats*, *two of *no cats*, *two of *more male than female cats*

Gen 7 *Post-*of *DPs of the form Det + Noun denote proper principal filters (= for some $p > 0$, $F(q) = True$ iff $p \leq q$).*

query Which NPs occur naturally in Existential There (ET) contexts, as in:
Aren't there *at most four undergraduate students* in your logic class
Weren't there *more students than teachers arrested* at the demonstration?
Just *how many students* were there at the party?
Aren't there *as many male as female students* in the class?
*There are *most students* in my logic class
*Isn't there *the student who objects to that*?
*Isn't there *every student who gave a talk at the conference*?
*Was there *neither student* arrested at the demonstration?

Gen 8 *Just NPs built from intersective Dets and their boolean compounds (modulo pragmatic factors) occur in ET contexts.*

Intersective (Generalized Existential) Dets are ones whose values at a pair A,B of properties just depends on $A \cap B$. Formally, they satisfy (3):

(3) $D(A)(B) = D(X)(Y)$ whenever $A \cap B = X \cap Y$.

some intersective Dets some, a/an, no, several, more than six, at least / exactly / fewer than / at most six, between six and ten, just finitely many, infinitely many, about / nearly / approximately a hundred, a couple of dozen, practically no, not more than ten, at least two and not more than ten, either fewer than five or else more than twenty, that many, How many?, Which?, more male than female, just as many male as female, no. . . but John

Co-intersective Dets *every*, *all but two*,. . . which satisfy (3) with $-$ for $\cap$, are not intersective. Nor are **proportionality Dets**: most, less than half, seven out of ten

3. Properties of PL not present in NL

Precision NL, not PL, is structurally ambiguous

1. John didn't leave because the children were crying
   R1: That's why he stayed [not leave][because the children were crying]
   R2. He left for some other reason [not [leave because the children were crying]]
   Compare in PL: $\neg(P \rightarrow Q)$ versus $(\neg P \rightarrow Q)$

2. Every student read a Shakespeare play (over the vacation)
   R1: For every student there was a play he read—maybe different students reaD different plays
   R2: There was one Sh. play that every student read (maybe Hamlet, maybe Lear,...)
   Compare in PL: $\forall x \exists y (x < y)$ vs $\exists y \forall x (x < y)$ They have different truth values

3. John told Bill that he had the flu. John said: "I have the flu", "You have the flu", or Henry (identified in context) has the flu. Compare: john$_x$(x told bill that x had the flu), bill$_y$(john told y that y had the flu), john told bill that z had the flu.

4. John thinks he's clever and so does Bill [think that John is clever, think that he himself is clever]
   John$_x$(x think x is clever & Bill think that x is clever)
   John$_x$(x think x is clever) & Bill$_y$(y think that y is clever)

Fact: NL lacks the variable binding operators of PL.

4. Logical resources of NLs not present in PL

NL quantifiers take pairs of properties as arguments, the first restricting the domain of quantification, as in *Most **poets** daydream*. PL quantifiers have just one property argument:

a. Some poets daydream = $\exists x(P(x)\&D(x)) \equiv SOME(\lambda x(P(x)\&D(x))$
b. All poets daydream = $\forall x(P(x) \rightarrow D(x)) \equiv ALL(\lambda x(P(x) \rightarrow D(x))$

Theorem (Keenan 1992) The domain eliminable NL quantifiers are just the (co)-intersective ones, thus excluding the proportionality Dets.

Gen 9 *All PL quantifiers are domain reducible; not so in NL.*

Def If Det is proportional then the truth of *Det poets daydream* depends on the proportion of poets that daydream. ($DAB = DXY$ whenever $|A \cap B| \,/\, |A| = |X \cap Y| \,/\, |X|$)

Examples: most, seven out of ten, less than half, not one... in ten
Most poets daydream does not mean either (For most objects x (Poet(x) & Daydream(x)) or (For most objects x, if Poet(x) then Daydream(x)). BUT

Gen 10 *a. NL Quantifiers are domain independent: Blik defined by* $BLIK(A)(B) =$ *T iff* $|\neg A| = 2$ *is not a possible English determiner.* Blik cats are black *would be true iff the number of non-cats is two.*

*b. NL Qs are overwhelmingly **conservative**: Det As are Bs cannot depend on Bs which are not As, so* $DAB = D(A)(A \cap B)$ *NB: Conservativity (CONS) and Domain Independence (DI) are independent. (BLIK is CONS but not DI; F in $FAB =$ T iff $|A| = |B|$ is DI but not CONS)*

Gen 11 *Proportionality Quantifiers determine novel reasoning paradigms:*
Exactly half the students passed. Therefore, Exactly half the students didn't pass.
Between a third and two thirds of the students passed the exam. Therefore, between a third and two thirds of the students didn't pass the exam.
$Qx\phi$ never entails $Qx\neg\phi$, for $Q =$'all' or 'some'

Gen 12 *Non-trivial Proportionality quantifiers are "logical" (= their denotations are permutation invariant) but **not definable** in PL. Similarly with cardinal comparatives, of type* $((1,1),1)$*:*
***More** poets **than** priests daydream; **Fewer** boys **than** girls, **More than twice as many** girls **as** boys; **Half again as many** girls **as** boys. These quantifiers may have multiple occurrences: **Fewer** boys **than** girls read **more** poems **than** plays. Jack read **more poems than Jill. A certain number of** students applied for **a smaller number** of scholarships.*

Gen 13 *PL quantifiers are **extensional**, NL ones may not be. In a situation in which the doctors and the lawyers are the same individuals,* Every doctor attended (the meeting) *and* every lawyer attended... *have the same truth value, but* Not enough doctors attended *and* not enough lawyers attended *may have different values. All PL quantifiers are like* every *here.* too many, surprisingly many, ... *are like* not enough.