International Journal of Soft Computing 13 (1): 31-36, 2018
ISSN: 1816-9503

Arithmetical Logic for AI Deep Learning

Yvon Gauthier
Faculty of Arts and Sciences, University of Montreal, Montreal, Quebee, Canada

Abstract: A logico-mathematical foundation for deep learning in artificial intelligence is proposed using Kronecker's theory of homogeneous polynomials and Hensel's lifting lemma in modular arithmetic. It is suggested that such a foundation is appropriate for the multileveled architec tures of deep learning scenarios in artificial intelligence.

Key words: Polynomial modular logic, modular arithmetic, Kronecker's theory of forms, Hen-sel's lifting lemma, foundation, homogeneous

INTRODUCTION

Researcher propose in this study a logico-arithmetical framework for the architectures of AI in the deep learning approach for pattern recognition in vision and language among other objectives. Arithmetical logic or polynomial modular logic is the internal logic of arithmetic from a constructivist point of view. Internal logic is a logic or deduction of content as Hilbert called it "inhaltiches logisches Schliessen". Researcher use here, Kronecker's theory of forms (homogeneous poly-nomials) as by Gauthier (2015, 2017) along with Hensel's lemma in modular arithmetic to delve into the depths of AI networks in supervised, semi-supervised or unsupervised training.

MATERIALS AND METHODS

Logic: In his excellent survey, Bengio (2009) to logical gates as propositional tautologies in a Boolean algebra for the connectives and $\wedge$ or $\vee$ and not $\neg$. He also mentions the pioneering research of McCulloch and Pitts (1943) who devised a logical calculus for artificial neurons as they say in the logical syntax of carnap and the formal language of russellian legacy. Logical gates and logical circuits are transposed on the electrical switches on/off in a bivalent Boolean algebra for truth values 1 or T (True) and 0 or F (False). Such a logic is adequate for classical logical systems and standard computer programming but it is well known that classical logic is not sufficient to formalize inferences in many mathematical theories, e.g., topos theory and other logics like intuitionistic logic are called upon oftentimes to rescue the content of logical deduction. However, logic in its different clothes, classical intuitionistic, multivalued, modal, relevant or linear, paraconsistent (with default reasoning) or even fuzzy logic does not seem to be able to capture the inner structure of human or machine learning. Functional languages like lambda calculus and type-theoretic methods or category-theoretic methods, even in their constructive versions are too close to the cumulative ordinal ranks of transfinite set theory to provide a fine structure for the layered organization of deep learning. On the other side, algorithmic complexity escapes the reach of classical methods that are bounded by in completeness results in Peano arithmetic and computations have a limited scope in NP complexity class in polynomial time.

Instead of logical investigations, research in artificial intelligence has concentrated on ap-proximation methods, iteration processes, regression analysis, large datasets mining, statistis-tical methods and stochastic or probabilistic models that provide approximate valuations or calculi for AI multileveled deep training and deep belief nets at the junction of artificial or mechanical intelligence, natural or biological intelligence (neuroscience) and cognitive science (human intelligence).

Arithmetic: Researcher claim in the following that arithmetical congruences in algebraic number theory provide an appropriate foundation for the nonlinear processes of AIDL and especially, the ring Z/pZ of integers modulo p a prime number will serve as the proper arena for dealing with nonlinear approximations. Algebraic integers are part and parcel of the theory of forms or homogeneous polynomials of higher degree developed by Kronecker and pursued by his student and edi-tor kurt hensel in his new foundation of algebraic numbers-which he says to be totally independent of the (Dedekindian) theory of ideals (Hensel, 1897).

It is Hensel who introduced p-adic arithmetic and what is called Hensel's lemma or Hensel's lifting lemma allows to lift iteratively the root of a polynomial modulo a prime power to higher powers in a finite ascent much in

the same way that Kronecker constructed his theory of forms or polynomials of higher degree as a linear combination of lower powers effected by a finite descent to linear polynomials which can be considered as first approximation linear polynomial congruences in modular arithmetic. Let us, review first Kronecker's theory as by Kronecker.

Kronecker wanted to construct a general arithmetic "allgemeine arithmetik" of algebraic quantities. Kronecker went on to elaborate a theory of the content (Inhalt or Enthalten-Sein in German) of homogeneous polynomials, divisor theory (Modulsyteme) and elimination theory which researcher briefly expose.

The general theory of elimination for polynomial equations proceeds along the lines of a general arithmetic of rational functions with integer coefficients and indeterminates. Forms (polynomials) can contain <enthalten> other forms or be contained in other forms and two forms are said to be "absolutely equivalent" when they contain each other. Definitions of primitive, prime, irreducible forms follow. It is useful to quote in full proposition 9th (Kronecker, 1968).

When a homogeneous linear form F is contained in another form F_0 the latter can be transformed in the former provided that forms of the domain (of rationality) are substitued for the indeterminates of F_0; those forms are linear if F_0 is itself a linear form. In such a case the contained linear form F is transformed into the containing form through a linear substitution with integral coefficients and this a sufficient condition for the containment <Enthalten-Sein> of F in F_0.

Kronecker explains that the linear substitution refers to the indeterminates and the integer coefficients are the entire rational functions or integral quantities of the domains of rationality R, R', R". Proposition X then ensues.

Equivalent homogeneous linear forms can be transformed one into the other through substitution with integral coefficients.

Divisibility properties are easily deducible, e.g., absolutely equivalent forms have the same divisors and the final conclusion is reached with the statement 13 (and 13°) on the unique factorization of integral algebraic forms as products of irreducible (prime) forms. What this shows, Kronecker maintains is that the fundamental laws of ordinary arithmetic are preserved in the encompassing sphere of algebraic quantities by the process of association of algebraic forms.

Kronecker's version of unique decomposition rests on the equation:

$$\prod_{h=1}^{r} M_k U_{hk}$$

and:

$$\prod_{i=j+k} c_i = \sum_{j+k=i} a_j b_k$$

Where:

M's = Integral forms

U's = Indeterminates

c's = Integral coefficients with j = (0, ..., m) and k = (0, ..., n)

We shall read it in the form (remembering that $a^{p-1} \equiv 1 (\text{mod } p)$ from a divisibility point of view):

$$\prod_{i=1}^{m+n}(1+c_i x_i) = \sum_{i=0}^{m+n}(c_i x^{m+n-1}) = \sum_{m+n=1}(a_m b_n)$$

Kronecker's generalization uses the convolution product for polynomials:

$$\sum_h M_h U_h \cdot \sum_i M_{m+i} U^{i-1} = \sum_k M_k' U^k$$

So that, the product:

$$\prod_h \sum_k M_k' U_{hk}$$

is "contained" in the resulting form and the product can be expressed as:

$$\sum_k M_k' U^k = (M_k M_{m+1})^k = (M_k M_{m+1})^{k-1} + (M_k M_{m+1})^{k-2} +, ..., +(M_k M_{m+1})$$

In the decreasing order of the rank k of the polynomial sum. This linear combination obtained by the convolution product and the finite descent of powers shows simply that integral rational forms generate integral algebraic forms, i.e., algebraic integers.

The outcome of Kronecker's construction for the content of modular arithmetic can be summarized in the equation:

$$M \equiv M' (\text{modd } M_1^0, M_2^0, M_3^0)$$

Meaning that, the difference M-M′ contains the modular systems, $M_1^0, ..., M_3^0$. Remark that in Z/pZ, for primes a and b, $a \equiv b \ (\text{mod } p^n)$ gives the difference $a-b = p^{n-1}$ which means that p^{n-1} is contained in p^n. Hensel's lemma simply states that if the polynomial $f(x) \in Z/pZ$ has a (simple) root $r \in Z/pZ$ which satisfies:

$$f(r) \equiv 0 \text{ mod } p, f'(r) \equiv 0 \text{ mod } p$$

where, f' (r) is the arithmetic derivative $(r^n)' = nr'^{n-1}$, then there is a unique $s \in Z/pZ$ such that:

$$f(s) \equiv 0 \text{ mod } p^{k+m} \text{ and } r \equiv s$$

for positive integers k, m with $k \leq m$. One can now lift a root r of the polynomial f (x) mod p^k to the new roots mod p^{k+1} to the effect that the roots of mod p^k are lifted to mod p^{k+1}. For p-adic integers, one can write for the roots r and s the equation:

$$s = r - \frac{f(r)}{f'(r)}$$

And the roots mod p^k or higher powers of p are better and better approximations to a unique root. If we look at polynomials with integer coefficients a and in determinates x:

$$f(x) = a_n x^n + a_{n-1} x^{n+1} +, ..., + a_2 x^2 + a_1 x + a_0$$

As finite formal power series with decreasing powers, we can see the indeterminate x becoming smaller and smaller meaning that it is the same modular process for the composition and the decomposition of approximations.

Hensel's lifting lemma is analogous to Newton's geometric (polygon) method for locating the real root of a differentiable function but the algebraic approach with modular arithmetic has the advantage of being constructive for it allows for a simple finite nonlinear calculus. The iterative process on powers constitutes progressively finer approximations, since, the initial congruence at level k = 1 is a linear approximation (of degee 1). The nonlinear ladder, quadratic k = 2 and higher degree polynomials in Kronecker's study of algebraic integers can be ascended for the composition of congruences as in Hensel's lemma or descended in the decomposition of powers as in Kronecker's theory of forms. In both cases, a finite process of construction is of the essence.

RESULTS AND DISCUSSION

Modular logic: Researcher briefly sketch here the modular polynomial translation of standard formal logic. The connectives negation, disjunction and conjunction are directly eliminable in the polynomial representation, since, one can interpret then as difference, sum and product of polynomials with a finite numbers of terms (coefficients and indeterminates).

Proof: We rewrite the logical intelim rules in the polynomial language with the quantifier ± meaning "for an unlimited sequence". The unique identity axiom becomes the equality A = A:

$$(I\wedge)\ \frac{A\ B}{A \wedge B};\ a_0x,\ b_0x \equiv a_0x.b_0x$$

$$(E\wedge)\ \frac{A \wedge B}{A} \text{ and } \frac{A \wedge B}{B};\ a_0x.b_0x \equiv a_0x, b_0x$$

$$(I\vee)\ \frac{A}{A \vee B} \text{ and } \frac{B}{A \vee B};\ a_0x+b_0x \equiv a_0x,\ a_0x+b_0x \equiv b_0x$$

$$(E\vee)\ \frac{\begin{matrix}A \vee B & [A] & [B] \\ & \vdots & \vdots \\ & C & C\end{matrix}}{C};\ \begin{matrix}a_0x+b_0x \equiv c_0x \pmod{b_0x} \\ a_0x+b_0x \equiv c_0x \pmod{b_0x}\end{matrix}$$

$$(I\rightarrow)\ \frac{\begin{matrix}[A] \\ \vdots \\ B\end{matrix}}{A \rightarrow B};\ a_0x \equiv b_0x \pmod{a_0x+1}$$

$$(E\rightarrow)\ \frac{A, A \rightarrow B}{B};\ 1\text{-}a_0x \equiv b_0x \pmod{a_0x}$$

$$(I\neg)\ \frac{\begin{matrix}A \\ \bot\end{matrix}}{\neg A};\ 1\text{-}a_0x \equiv b_0x\ 1 \pmod{a_0x}$$

$$(E\neg)\ \frac{A, \neg A}{\bot};\ 1\text{-}a_0x \equiv 0 \pmod{a_0x}$$

$$(I\forall)\ \frac{Ax}{\forall x\, Ax};\ \prod_n a_0x^n \equiv a_0x \pmod{n}$$

$$(E\forall)\ \frac{\forall x Ax}{At};\ a_0x \equiv \prod_n a_0x^n \pmod{1}$$

$$(I\exists)\ \frac{At}{\exists x\, Ax};\ \sum_n a_0x \equiv a_0x^n \pmod{1}$$

$$(E\exists)\ \frac{\begin{matrix}\exists xA & [Ax] \\ & \vdots \\ & B\end{matrix}}{B};\ a_0x \equiv \sum_n box^n \pmod{1}$$

$$(I\overline{\pm})\ \frac{Ax_n}{\overline{\pm} xAx};\ \prod_{n...} a_0x \equiv a_0x^n \pmod{n \times n}$$

$$(E\overline{\pm})\ \frac{\overline{\pm} xAx}{At_0};\ a_0x \equiv \prod_{n...} a_0x^n \pmod{1}$$

In translating logical formulas into congruent forms, we want to represent logical constants in a polynomial language in order to integrally arithmetize (polynomialize) logic. It is manifest in that context that deduction expressed in a turnstile A⊢ A or A/A is a congruence relation in a modular calculus. Implication is rewritten:

$$(\bar{a}_0x+b_0x)^n$$

for $\bar{a}_0x = 1-a_0x$, the local negation (complement) of logic; exponent n denotes the degree of the polynomial (content) of implication that, we reduce in the following way by a calculus on symmetrical polynomials (forms).

We want to arithmetize (local) implication. We put 1-a = $\bar{a}$ for local negation. We have $(\bar{a}_0x+b_0x)^n)$ and we want to exhaust the content of implication-in gentzenian terms, this would correspond to the exhibition of subformulas (the subformula property). We just expand the binomial by decreasing powers:

$$(\bar{a}_0x+b_0x)^n) = \bar{a}_0^nx+na^{n-1}x+[n\frac{(n-1)}{2!}]\,a^{n-2}xb^2x+,\ldots,+b_0^nx$$

where the companion indeterminate x shares the same power expansion. By an arithmetical calculation (on homogeneous polynomials that are symmetric, i.e. with a symmetric function f(x, y) = f(y, x) of the coefficients):

$$\begin{aligned}(\bar{a}_0x+b_0x)^n) &= \bar{a}_0^nx+\sum_{k=1}^{n-1}(\frac{n-1}{k=1})\,\bar{a}_0^{k-1}x+(\frac{n-1}{k})\,a_0^kxb_0^{n-k}x+b_0^nx\\ &= \sum_{k=1}^{n}(\frac{n}{k-1})\,a_0^kxb_0^{n-k}x+\sum_{k=0}^{n-1}(\frac{n-1}{k})\,a_0^kxb_0^{n-k}x\\ &= \sum_{k=0}^{n-1}(\frac{n-1}{k})\;a_0^{k+1}xb_0^{n-k}x+\sum_{k=0}^{n-1}(\frac{n-1}{k})\,a_0^kxb_0^{n-k}x\\ &= \bar{a}_0\sum_{k=0}^{n-1}(\frac{n-1}{k})\,(\bar{a}_0-1)^k\;b^{n-1-k}x+\sum_{k=0}^{n-1}(\frac{n-1}{k})\,\bar{a}_0^kx\,(b_0-1)^{n-1-k}x\\ &= (\bar{a}_1x+b_1x)\,(a_1x+b_1x-1)^{n-1}\end{aligned}$$

And continuing by descent and omitting the x's, we have:

$$\begin{array}{l}(\bar{a}_2+b_2)\,(\bar{a}_2+b_2-2)^{n-2}\\ \quad\vdots\quad\vdots\quad\vdots\quad\vdots\\ (\bar{a}_{n-2}+b_{n-2}+\bar{a}_{n-2}+_{n-2}-(n-2))^{(n-(n-2))}\\ (\bar{a}_{n-1}+b_{n-1}+\bar{a}_{n-1}+_{n-1}-(n-1))^{(n-(n-1))}\\ (\bar{a}_n+b_n)\,(\bar{a}_n+b_n)^{n-n}\end{array}$$

Applying descent again on $(\bar{a}_n+b_n)$, we obtain:

$$(\bar{a}_0+b_0)$$

or reinstating the x's:

$$(\bar{a}_0x+b_0x)$$

Remembering that:

$$(\bar{a}_x+b_x)^n_{k<n} = \sum_{k+m=n}(\frac{k+m}{k})\,\bar{a}^kb^mx^n$$

we have:

$$(\bar{a}_x+b_x)^{n+m=n}_{k<n} = \prod_{k+m=n}(k,m)\,2^n$$

or more explicitly:

$$\sum_{i\,=\,1}^{m+n}c_ix^{m+n=1} = \bar{a}_0x\times b_0x\prod_{i=1}^{m+n}(1+c_ix) = 2^n$$

where the product is over the coefficients (with indeterminates) of convolution of the two polynomials (monomials) a_0 and b_0. We could of course calculate the generalized formula for polynomials:

$$(a_0x+b_0x+c_0x+,\ldots,+k_0x)^n = \sum_{p,\,q,\,r\ldots s}a^pb^qc^r,\ldots,k^s$$

in the same manner. The combinatorial content of the polynomial is expressed by the power set 2^n of the n coefficients of the binomial. Researcher contend that this combinatorial content expresses also the meaning of local (iterated) implication. Convolution exhibits the arithmetic connectedness that serves to render the logical relation of implication. Implication is seen here as a power of polynomials, a^k and b^m with k<m having their powers summed up and expanded in the binomial expansion.

Some other formula may be used for the product but it is essential to the constructive inter-pretation that the arithmetic universe be bounded by 2^n. One way to make things concrete is to analyse a→b in terms of:

$$a \rightarrow b = C((2^n-a)+b)$$

where, C can stand for combinations or coefficients. The equation is an arithmetical analogue of the topological interpretation of intuitionistic implication.

- Theorem; Local implication a→b can be eliminated by interpreting it as $(\bar{a}+b)^n$
- Proof; By the above construction

A nice example of this polynomial translation is the probability calculus with binomialn probabilities and

conditional and inverse probability (Baye's theorem) as it is used in DL for the so, called Boltzmann energy machines. Bengio (2009) notes the limitations of Boolean propositional logic with logic gates and circuits and their associated algorithmic complexity. He points out the import of second-degree-rather than second-order in the text-polynomials for nonlinear processes in Boltzmann machines or energy-based machines models for deep belief networks. The so-called restricted Bolzmann machines exhibit a nonlinear behavior as they discharge energy:

$$E_i = mv^2$$

With symmetric probability distributions P(h/x) and P(x/h) for the energy function e(x, h) where x is the observed part and h the hidden part of the units of the machine environment with conditional probabilities:

$$P\left(\frac{h}{x}\right) = \prod P\left(\frac{h_i}{x}\right)$$

and:

$$P\left(\frac{x}{h}\right) = \prod P\left(\frac{x_i}{h}\right)$$

This formulation can be given an easy polynomial translation in arithmetical logic.

The fundamental arithmetic congruence relation is basic to all logical calculi from combi-natorial logic to algebraic logic and categorical logic including the probability calculus. We have:

$$P\left(\frac{A}{B}\right) = P(AB)P(B)$$

the conditional probability of A given B gives:

$$P(\bar{a}_0x+b_0x) = \frac{P(\bar{a}_0x\times b_0x)}{P(\bar{a}_0x)}$$

and Bayes theorem stating:

$$P\left(\frac{A}{B}\right) = \frac{P(B/A)\,P(A)}{P(B)}$$

gives:

$$P(\bar{a}_0x+b_0x) = \frac{P(b_0x+\bar{a}_0x)P(\bar{a}_0x)}{P(b_0x)}$$

and for the inverse probabibily:

$$P(E_i/A) = \frac{P(E_i)P(A/E_i)}{\sum_{j=1}^{n} P(E_j)P(A/E_j)}$$

We have:

$$P(e_ix+\bar{a}_0x) = \frac{P(e_ix)P(\bar{a}_0x/e_ix)}{\sum_{j=1}^{n} P(e_jx)P(\bar{a}_0x/e_jx)}$$

CONCLUSION

In this theoretical study, researcher wished to provide a foundational logico-mathematical justification of AIDL within the conceptual framework of constructive arithmetical logic. LeCun *et al.* (2015) member of the pioneering trio of researchers in AIDL with Bengio (2009) has evoked the influence of Jean Piaget's constructivist (genetic) epistemology of learning in the sensory-motor development of intelligence. Although, piaget had insisted on constructivist motives in his child psychology his logic of operations (logique operatoire) with the Klein four-group or Boolean group INRC is too shallow a propositional basis to be regarded of foundational import for AIDL. Bengio (2009) has emphasized the limitations of propositional Boolean logic (a linear logic among others) and the limits of algorithmic complexity while Hinton has highlighted the nonlinear nature of deep belief networks. Those concerns are formally taken into account in the internal logic of arithmetic, the polynomial modular logic outlined in this study. Researcher proposal in a nutshell could be considered as a deep (discrete contra continuous) approximation theory in accord with the finite multilayered struc-ture of AIDL. The algebraic-arithmetical theory of congruences with its Koneckerian ancestry can serve as a foundational background for computer-theoretical research in DL. For instance, the equivalence theorem by Montufar and Morton (2015) on the content of a mixture of products in a product of mixtures for exponentially large number of parameters in probability distributions is reminiscent of Kronecker's idea of the content of polynomials with integer coefficients Kronecker's equivalence principle says the content of the product of two polynomials is the product of their respective contents. The multileveled machinery of deep learning may require an ascent up to 110° polynomials as Montufar and Morton show of course, researcher do not claim any concrete result, nor do researcher suggest any implementation procedure. The purely syntactic scheme researcher have proposed is not intended as a model-theoretic superstructure of AIDL but as a constructivist logico-mathematical infrastructure for the architectures of AIDL (Gauthier, 2013, 2015).

REFERENCES

Bengio, Y., 2009. Learning deep architectures for AI. Found. Trends Mach. Learn., 2: 1-127.

Gauthier, Y., 2013. Kronecker in contemporary mathematics general arithmetic as a foundational programme. Reports Math. Logic, 48: 37-65.

Gauthier, Y., 2015. Towards an Arithmetical Logic: The Arithmetical Foundations of Logic. Springer, Berlin, Germany, ISBN:978-3-319-22086-4, Pages: 179.

Hensel, K., 1897. On a new grounding of the theory of algebraic numbers. Annu. Rep. Ger. Assoc. Math., 6: 83-88.

LeCun, Y., Y. Bengio and G. Hinton, 2015. Deep learning. Nature, 521: 436-444.

McCulloch, W.S. and W. Pitts, 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys., 5: 115-133.

Montufar, G.F. and J. Morton, 2015. When does a mixture of products contain a product of mixtures?. SIAM. J. Discrete Math., 29: 321-347.