ISSN: 2641-2039
Global Journal of Engineering Sciences
DOI: 10.33552/GJES.2023.11.000753
Iris Publishers

Mini Review

The Algebra of Artificial Intelligence

Arvid Siqveland*

Institute of Natural Science and Industrial Systems, University of South-Eastern Norway, Norway

***Corresponding author:** Arvid Siqveland, Institute of Natural Science and Industrial Systems, University of South-Eastern Norway, Norway.

Received Date: March 06, 2023

Published Date: March 14, 2023

Abstract

Machine learning engineering is most commonly about determining the coefficients in experimentally constructed models. From quantum mechanics, we learn that the models themselves are determined by the intrinsic geometry of the dynamical system under study. It is possible to use Turing Machines and mathematical logic to study machine learning, which is what should be called Artificial Intelligence. Finally, Probabilistic Turing Machines and distributions can be studied by algebraic geometry.

Motivation

I have used Google to learn about Artificial Intelligence (AI). I have read Dan Brown's "Origin", and I even found a book on AI in pdf format [1], that can be used as a textbook in an undergraduate module on the subject. Also, I supervised a bachelors project on machine learning (ML), recognizing litter, in the spring of 2018.

I have to admit that I studied Informatics and at the University of Oslo up to a bachelor's degree. During that study, I programmed several self-learning systems, where the most funny and useless was a "German language generator", learning German from reading 5000 pages of Goethe and then using a random algorithm to produce a German novel.

So, to be honest and humble, this is what I know about the subject, and this lecture is about one possible contribution from contemporary mathematics.

Introductory Definition of AI

A machine X is a state of parameters that produces the next state. This can be given in diagram form, and any algorithmic procedure can be defined by a finite state machine. If there is a converging state, the machine is called a Turing machine after Alan Turing, and Gödel proved that not all mathematical processes can be defined by a Turing Machine: That is: Some processes need intelligence beyond computers.

A self-learning machine is a state of parameters that produces the next state based on all previous states. That is why we say that the machine has learned how to act.

Definition 1: An intelligent machine is a state of parameters

$$\wedge_m = \left(\lambda_1(m),\ldots,\lambda_n(m)\right) \in \mathrm{M}^n$$

such that the next state is

$$\left(X_1\left(\wedge_{\leq m}\right),\ldots,X_n\left(\wedge_{\leq m}\right)\right)$$

where $X_1, \ldots, X_n$ are stochastic regression variables on the universe of all previous states and M is the moduli (the classifying space) of the parameters.

Conclusion: For me, *AI* is the mathematical study of *AI* machines.

Thinking Inside Boxes

Mathematical statistics

At first, because there are stochastic regression variables involved, one looks into the toolbox of statistics. One finds that the machines we can study with ordinary statistics are the machines with real parameters. To be very rude, undergraduate statistics studies the probability distribution of a stochastic variable *X* with probability distribution (function)

$$f_X : \mathbb{R} \to \mathbb{R}$$

and can be applied to single-valued numerical machines. Then the probability of the next state is

$$P(a \le X \le b) = \int_a^b f(x)dx.$$

You engineers should notice that when constructing the machine, the next state is defined as the expected state *E(X)* of the stochastic variable *X*.

Digging deeper in the statistical toolbox

Most interesting machines have more than one parameter (The only one parameter model that comes to me at the moment is "What is your next weight if you continue eating like this"). So, we need to consider probability distributions that are functions

$$f : \mathbb{R}^n \to \mathbb{R}^n$$

leading to probability distributions given by regions $R \subseteq \mathbb{R}^n$:

$$P(X \in R) = \int_R f(x_1, \ldots, x_n)dR.$$

This is where the statisticians understand that they need the toolbox called multi-variable calculus, or differential geometry. Also, what if the parameters (variables) are not real numbers, but some other mathematical objects? That is to say, the tools in the box fit other units.

Mathematical Models Workshop

We refer to the book [2]. A mathematical model of a natural phenomenon *P* is a mathematical object *X* depending on parameters such that the changing aspects (states) of *P* correspond to altered parameter-values for *X*, and such that any choice of parameter-values corresponds to some possibly occurring aspect (state) of *P*. This is to say that there is a surjective map of sets

$$X : \{states\ of\ \mathrm{P}\} = \mathrm{P}^M \to \chi\ ;$$

where *X* is the set of possible models, and the states of the phenomenon *P* is given by parameters in *P*. Notice that the parameters might be indexed by an infinite set *M*.

Calibrating the models

Different parameter values might correspond to equal states: Saying that we have some parameters that are redundant for some states. We say that two models are equivalent, $X(1) \sim X(2)$ if they correspond to the same object.

This says that two sets of parameters should be identified if they correspond to equal states, so that a mathematical model is a representation of $P^M/\sim$, that is a map of sets

$$X : (\mathrm{P}^M / \sim) \to \chi :$$

A Possible Mathematical Model of AI

Definition 2: *A mathematical model of an intelligent machine (IM) is a mathematical object X(IM) in the state space P/ ~ of the machine.*

The change of states of the machine is given by the geometry of the space when we identify geometry with the possible dynamics on the moduli.

Recall: Short and Heuristically

A finite state machine is a machine that can be in a finite number of states. Its next state is determined by its current state.

An intelligent (self-learning) machine, is a machine that can be in an infinite number of states, and its next state is determined by all previous states.

Notice that if we define a distance (metric) on the state space, then we can compute the rate of change per distance, and by that define the concept of time.

We hereby indicate that a model of the universe as a self-learning machine might work, causing problems in most religions.

The ancient toolbox of differential geometry

This is in fact the generalization of the study of continuous functions

$$f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}^m$$

for positive numbers *n,m*. Of course, the concepts of a topolgy (open sets) and thereby continuity is a rather advanced concept by itself. The homomorphism *f* defines a dynamic on its image (is graph) by usual partial derivation (in fact, without involving any time until we parametrize the graphs by a time *t*).

The modernized toolbox of differential geometry

The generalization consists in covering any geometric figure (topological space, as the hemisphere) called a manifold by open charts isomorphic to Rn, and then defining dynamic properties of functions

$$f : M_1 \to M_2$$

between two manifolds, by giving them local properties that can be glued in the intersection of the charts.

Of course, to solve some explicitly given differential equations, we are in the need of (at least) the complex numbers, so a near to final state of differential geometry is the theory of complex manifolds, replacing $\mathbb{R}^n$ by $\mathbb{C}^n$.

The fundamental theorem of complex manifolds

Theorem 1: Any analytic (infinitely many times differentiable) function $f:\mathbb{C}^n \to \mathbb{R}^n$ is given by a uniformly convergent power-series,

$$f(z_1,\ldots,z_n)=\sum_{i=1}^{\infty} a_{i\underline{z}^{n_i}}$$

where $\underline{z}^{n_i} = z_1^{n_1} z_2^{n_2} \ldots z_n^{n_n}$ is a monomial in $z_1, \ldots, z_n$, of degree $i = n_1 + \ldots + n_n$.

Humans are self-learning machines?

The following is not something I claim, I just state it as a possible model: Our behaviour is based on our history, and the teacher just ads previous history to our state. Performing algorithmic behaviour are instincts or programmed behaviour, not intelligent behaviour. Because of this, I would like to tell you some of the history of algebra. Then the state consisting of knowledge of algebra, differential geometry and computer programming leads to the study of algebraic geometry as a tool for artificial intelligence.

Algebraic Geometry

There is no clear history of algebra. For Norwegian citizens, we like to say that it started with Niels Henrik Abel from Gjerstad (1802 - 1829). In short, we can say that when the Italians struggled to give a formula for the solutions of the equation

$$x^5 + ax^4 + bx^3 + cx^2 + dx + e = 0$$

that is a formula for the roots of the polynomial, Abel showed that the sets of roots of any polynomial has a particular structure making the set which is now called a group, or a root field. This says that there is no general "ordinary" formula for the roots of a 5th order polynomial.

The last century of algebraic geometry

The last 300 years or so, algebraic geometry is used as the technology of going from geometric figures in Rn to algebraic structures. The ultimate goal was to prove Fermat's last theorem:

Theorem 2: There is no integer solutions to the equation

$$x^n + y^n - z^n = 0$$

for n > 2.

Finally, after an intense working period from he was about 7 until he was 43, Andrew Wiles proved an algebraic equivalent of Fermat's last theorem by the geometric theory of hyperelliptic curves. (At UiO, there is no longer a group in algebraic geometry: The old group is merged with topology and is renamed Algebra and Geometry).

The algebraic geometric toolbox

The new generation mathematicians immediately see that when they start programming differential geometry, as computers are finite state machines, any function is approximated by a polynomial (rather than a power series). I don't know who's come up with the idea (I do, it was Oscar Zariski), but we make a simplified differential geometry, by taking our spaces to be covered by Cn, and our functions

$$f:\mathbb{C}^n \to \mathbb{C}^m$$

to be only the rational functions $f(z_1,\ldots,z_n)=\frac{p(z_1,\ldots,z_n)}{q(z_1,\ldots,z_n)}$ where p and q are rational. This includes giving a topology on Cn consisting of the complements of the closed sets, which are the zero-sets of polynomials.

Dynamical Structures

"Suddenly" (we have yet no time, what is suddenly?), the study of the geometric structure is the study of the algebra of all polynomials in n variables, with some relations between the variables. This is what is a called a polynomial algebra

$$A = k\left[z_1, z_2, \ldots, z_n\right]$$

To give a dynamical structure on the geometry, is to give differential equations in the variables, and any set of relations between differential operators is what Sophus Lie (1842 - 1899) called a transformation group. This is now called a Lie algebra, and just recently, a dynamical structure is an action of a Lie algebra g on a polynomial algebra.

Associative algebraic geometry

Here, I refer to the book [3], which is again based on the book [4]. In this book I prove that the study of algebraic classifying spaces can be studied by the properties of a single algebra, the coordinate ring of the system. Thus

computer algebra can be directly applied to AI, and as the coordinate ring is noncommutative and in the general case very big, this is applying to the theory of Big Data. This says, we really need fast computers, maybe even quantum computers.

A Program for artificial intelligence

- Mathematical logic and the definition of a self-learning state machine
- II. Differential geometry and partial differential equations (PDE's)
- Category theory to prove that differential geometry over C is equivalent to complex algebraic geometry. Fun fact: John Nash proved that Navier-Stokes are solvable in algebraic geometry, so they are in diff. geom. also).

- IV. Noncommutative algebraic geometry by deformation theory to classify dynamical structures, and to prove that a global time can be defined, restricting locally to local time: Uniting General Relativity and Quantum field theory.

Acknowledgement

None.

Conflict of Interest

No conflict of interest.

References

1. Stuart Russel, Peter Norvig (2010) Artificial Intelligence, a Modern Approach, Third edition, Prentice Hall.
2. OA Laudal (2021) Mathematical models in science, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.
3. Arvid Siqveland (2023) Associative Algebraic Geometry. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.
4. E Eriksen, OA Laudal (2017) A Siqveland Noncommutative Deformation Theory.