

# Logic, Topos and Learning Theory

Xiyu Zhai

## Contents

<b>1</b>	<b>Statistical Axiom System</b>	<b>1</b>
<b>2</b>	<b>Related Concepts</b>	<b>2</b>
2.1	Expert System . . . . .	2
2.2	Knowledge Graph . . . . .	2
<b>3</b>	<b>Process of Falsification</b>	<b>2</b>
3.1	Two Axioms System . . . . .	2
3.2	N-Cyclic Axioms System . . . . .	3
3.3	Topological Elements . . . . .	3
<b>4</b>	<b>Topos</b>	<b>3</b>

This note promises a gold marriage between formal methods in computer science, statistics and abstract mathematical theories of logic and topos.

## 1 Statistical Axiom System

Deep learning is known to require a lot of training data to work, however human doesn't need that amount of data to perform well in his/her life.

The key, we believe, is to manage statistical complexity of a statistical axiom system.

A **statistical axiom system**, differs from an axiom system in that each axiom is only statistically believed to be correct.

A statistical axiom system is of **a dynamic nature**. New axioms can be added by observing facts, old axiom can be removed, either by directly using a factual counterexample, or by proving false via formal reasoning.

It needs **significant less samples than end-to-end approach** because it bases its correctness on a finite set of axioms. A prediction is derived from these axioms.

In a practical scenario of reinforcement learning or language processing, we constantly need to tell something is true or not. But we don't have a complete set of ground truths, and we also don't necessarily have enough samples.

## 2 Related Concepts

### 2.1 Expert System

Expert system lacks the ability to learn from data.

### 2.2 Knowledge Graph

Knowledge graph can't reason formally, statistical complexity is too large, just like end-to-end approach.

## 3 Process of Falsification

A statistical axiom system might not be logically consistent. In fact, we often need to choose a maximally likely subset that is logically consistent, or equivalently, prune away false axioms.

**Remark** (Origin of Mathematics). *This is a philosophical discussion of how mathematics is generated.*

*The belief is that mathematics doesn't come from a priori knowledge embedded in our brain.*

*Mathematics is basically knowledge generated by a set of axioms that have never been falsified in the entire history of human beings.*

In general, the problem of falsification is quite hard at first glance. So let's study simple examples first.

### 3.1 Two Axioms System

Thinks of the following scenario:

you have two friends A and B. A believes that Michael Jordan is the best basketball player ever, whereas B believes that LeBron James is the best basketball player ever. As there could be only one best basketball player ever, and Michael Jordan is a different person than LeBron James, only one of them could be correct.

Suppose that A and B each have an independent probability of being correct  $p_A > 1/2$  and  $p_B > 1/2$  respectively. There are only two logically consistent cases:

- A is correct and B is wrong, with probability  $p_A(1 - p_B)$
- B is correct and A is wrong, with probability  $p_B(1 - p_A)$

Then if  $p_A/(1 - p_A) > p_B/(1 - p_B)$ , we are more inclined to believe that A is correct and B is wrong.

### 3.2 N-Cyclic Axioms System

Suppose we have  $n$  people, the  $i$ th person believes that he's better than  $i - 1$ th person in mathematics (where  $i - 1$  is  $n - 1$  when  $i = 0$ ).

Suppose that  $i$ th person is correct with probability  $p_i > 1/2$ .

We use the word 'configuration' to mean assignments of each person with being correct or wrong.

The goal is to find a configuration with maximal likelihood.

Notice that each  $p_i > 1/2$  implies that such an optimal configuration must be maximal, in the sense that any configuration with extra corrects must be logically inconsistent.

It follows quite easily that any maximal logically consistent subset must be the whole set except just one point.

So the solution is then quite easy: find  $i$  such that  $p_i/(1 - p_i)$  is minimal and choose the  $i$ th person to be wrong, and everyone else should be correct.

**Remark.** *This problem is equivalent to the topological argument that there is no function on  $S^1 \rightarrow \mathbb{R}$  such that it gives identity  $id_{S^1}$  when composed with the canonical map  $\mathbb{R} \rightarrow S^1$ .*

**Remark.** *Cyclic Proofs are actual cycles in the space of axioms*

### 3.3 Topological Elements

Let's jump to topological formulation directly.

(At this point, I'm not sure how this compared with the mathematical theory of topos in details. But they're related closely at a higher level)

Let  $S$  be the set of all possible realities. An element  $s \in S$  contains every information in that reality.

Let  $X$  be a space of axioms. Each  $x \in X$  can be thought as a subset of  $S$  so that each  $s \in x$  means that  $s$  is a reality where  $x$  is correct.

For each  $s \in S$ , we define  $X(s) := \{x \in X : s \in x\}$ .

We define a topological structure over  $X$  to be the minimal one such that each  $X(s)$  is open.

**Example** (unit circle). Let  $S$  be the set of all continuous functions from  $S^1$  to  $\mathbb{R}$ .

Let  $X$  be indexed by  $S^1$ , with  $x_p := \{s \in S : e^{2\pi i s(p)} = p\}$ .

Then the topological generated is the same as induced by the identification of  $X$  with  $S^1$ .

## 4 Topos