CAS 701 Fall 2002

07 Practical Application of The Axiomatic Method

Instructor: W. M. Farmer

Revised: 28 November 2002

What is the Axiomatic Method?

1. A mathematical model is expressed as a set of axioms (in a language) called an **axiomatic theory**

2. New concepts are introduced by making **definitions**

3. Assertions about the model are stated as **theorems** and proved from the axioms

Notes:

- The axiomatic method is a method of **presentation**, not a method of **discovery** (Lakatos)

- The axiomatic method can be used as a method of **organization**

History

- Euclid (325–265 BC) used the axiomatic method to present the mathematics known in his time in the **Elements**
  - The axioms were considered truths
- The development of **noneuclidean geometry** by Bolyai, Gauss, and Lobachevskii (early 1800s) showed that axioms may be considered as just assumptions
- Whitehead and Russell formalized a major portion of mathematics in the **Principia Mathematica** (1910–1913)
- Bourbaki (mid 1900s) used the axiomatic method to codify mathematics in the 30 volume **Eléments de mathématique**
- Several libraries of formalized mathematics have been developed since the late 1980s using interactive theorem provers: HOL, IMPS, Isabelle, Mizar, Nqthm, Nuprl, PVS

Axiomatic Theories

- **Theory** = formal language + set of axioms
- **Language**: vocabulary for objects and their properties
  - Has a precise semantics (with a notion of logical consequence)
  - Can be used to describe multiple situations
  - The language usually belongs to a **logic**
- **Axioms**: assumptions about the objects and properties
  - Specify a class of **models**
  - Basis for proving **theorems**

Example: Theory of Partial Order

- Language: A first-order logic language with a 2-place predicate symbol $\leq$
  - $a \leq b$ is intended to mean a is less than or equal to b
- Axioms:
  - **Reflexivity**. $\forall x \, . \, x \leq x$
  - **Transitivity**. $\forall x, y, z \, . \, (x \leq y \land y \leq z) \supset x \leq z$
  - **Antisymmetry**. $\forall x, y \, . \, (x \leq y \land y \leq x) \supset x = y$
- The theory has infinitely many nonisomorphic models

Example: Peano Arithmetic

- Language: A second-order logic language with a constant symbol 0 and 1-place function symbol S
  - 0 is intended to represent the number zero
  - S is intended to represent the successor function, i.e., $S(a)$ means $a + 1$

- Axioms:
  - **0 has no predecessor**. $\forall x \,.\, \neg(0 = S(x))$
  - **S is injective**. $\forall x, y \,.\, S(x) = S(y) \supset x = y$
  - **Induction principle**.
    $\forall P \,.\, (P(0) \wedge \forall x \,.\, P(x) \supset P(S(x))) \supset \forall x \,.\, P(x)$

- Second-order Peano arithmetic is **categorical**, i.e, it has exactly one model up to isomorphism

Benefits of Axiomatic Theories

- **Conceptual clarity**: inessential details are omitted
- **Generality**: theorems hold in all models
- **Dual purpose**: a theory can be viewed as:
  - An abstract mathematical model
  - A specification of a collection of mathematical models

Theory Interpretations

- A **translation** Φ from T to T' is a function that maps the primitive symbols of T to expressions of T' satisfying certain syntactic conditions

- Φ determines:
  - A mapping of expressions of T to expressions of T'
  - Set of sentences called **obligations**

- Φ is an **interpretation** if it maps the theorems of T to theorems of T'
  - Sufficient condition: the obligations of Φ are theorems of T'

- **Interpretations are information conduits!**

Example: Theory of Computer Networks

- Theory name: Networks
- Language: Many-sorted first-order logic language with the following sorts and function symbols:

Sorts	**Function symbols**
boxes	box-of-interface
wires	wire-of-interface
interfaces	address-of-interface
addresses	

- Example axioms:
  - "Every box has a unique loopback interface"
  - "The address of a loopback interface is 127.0.0.1"

Example: Theory of Bipartite Graphs

- Theory name: Bipartite Graphs

- Language: Many-sorted first-order logic language with the following sorts and function symbols:

Sorts	**Operators**
red-nodes	red-node-of-edge
blue-nodes	blue-node-of-edge
edges	

- No explicit axioms

Example: Bipartite Graphs to Networks

- Let $\Phi_{\mathsf{BG} \rightarrow \mathsf{N}}$ be the translation from Bipartite Graphs to Networks defined by:

  red-nodes $\mapsto$ boxes

  blue-nodes $\mapsto$ wires

  edges $\mapsto$ interfaces

  red-node-of-edge $\mapsto$ box-of-interface

  blue-node-of-edge $\mapsto$ wire-of-interface

- $\Phi_{\mathsf{BG} \rightarrow \mathsf{N}}$ has no obligations

- $\Phi_{\mathsf{BG} \rightarrow \mathsf{N}}$ is an interpretation
  - "Transitivity of red-to-red connectivity" maps to "transitivity of box-to-box connectivity"

Example: Symmetry Interpretation

- Let $\Phi_{\mathsf{BG}\to\mathsf{BG}}$ be the translation from Bipartite Graphs to Bipartite Graphs defined by:

  red-nodes $\mapsto$ blue-nodes

  blue-nodes $\mapsto$ red-nodes

  edges $\mapsto$ edges

  red-node-of-edge $\mapsto$ blue-node-of-edge

  blue-node-of-edge $\mapsto$ red-node-of-edge

- $\Phi_{\mathsf{BG}\to\mathsf{BG}}$ has no obligations:

- $\Phi_{\mathsf{BG}\to\mathsf{BG}}$ is an interpretation
  - "Transitivity of red to red connectivity" maps to "transitivity of blue to blue connectivity"

Two Versions of the Axiomatic Method

1. **Big Theory**: A body of mathematics is entirely represented in one theory
   - Often a powerful, highly expressive theory like set theory is selected
   - All reasoning is performed within this single theory

2. **Little Theories**: A body of mathematics is represented as a network of theories
   - Bigger theories are composed of smaller theories
   - Theories are linked by interpretations
   - Reasoning is distributed over the network

Benefits of Little Theories

- Mathematics can be developed using the most appropriate vocabulary at the most appropriate level of abstraction

- Emphasizes reuse: if A is a theorem of T, then A may be reused in any "instance" of T

- Enables perspective switching

- Enables parallel development

- Inconsistency can be isolated: there are no interpretations of an inconsistent theory in a consistent theory, so inconsistency cannot spread from one theory to another

Formalized Mathematics

- **Mathematics** is a process of creating, exploring, and connecting mathematical models

- **Formalized mathematics** is the practical application of the axiomatic method within a formal logic
  - The mathematics process is performed with the aid of **mechanized mathematics systems**
  - Axiomatic theories are formally developed
    - Theory creation
    - Conservative theory extension
    - Theory exploration
    - Theory interpretation

Theory Creation

- Theories can be created in a several ways:
  - From scratch
  - By forming a union of a set of theories
  - By adding new vocabulary and axioms to a theory
  - By instantiating a parameterized theory
  - By instantiating a theory via an interpretation
- A theory may be required to contain a **kernel theory** which includes the machinery common to all theories

Conservative Theory Extension

- A conservative extension T' of T adds new machinery to T without compromising the original machinery of T

- The **obligation** of a purported conservative extension is a formula that implies that the extension is conservative

- Since T and T' are essentially the same theory, T' can be implemented by overwriting T
  - Avoids a proliferation of closely related theories

- There are two important kinds of conservative extensions that add new vocabulary to a theory:
  - Definitions
  - Profiles

Definitions

- A **definition** is a conservative extension that adds a new symbol s and a defining axiom $A(s)$ to a theory T
  - In some logics, the defining axiom can have the form $s = D$ (where s does not occur in D)

- The obligation of the definition is

$$\exists\, !x \;.\; A(x)$$

- The symbol s can usually be eliminated from any new expression of involving s

Profiles

- A **profile** is a conservative extension that adds a set $\{s_1, \ldots, s_n\}$ of symbols and a profiling axiom $A(s_1, \ldots, s_n)$ to a theory T

- The obligation of the profile is

$$\exists\, x_1, \ldots, x_n \,.\, A(x_1, \ldots, x_n)$$

- The symbols $s_1, \ldots, s_n$ cannot usually be eliminated from expressions involving $s_1, \ldots, s_n$

- Profiles can be used for introducing:
  - Underspecified objects
  - Recursively defined functions
  - Abstract datatypes

Theory Exploration

- The logical consequences of a theory are explored by:
  - Proving conjectures
  - Performing computations

- Products of theory exploration:
  - Theorems
  - Proofs
  - Counterexamples
  - Computations

- Tools of theory exploration:
  - Theorems
  - Transformers

Theorems

- Facts about a theory are recorded as theorems

- A theorem is usually installed in a theory only if it has been verified by a proof

- A theorem may sometimes be installed without a proof:
  - A theorem verified by a decision procedure
  - A theorem verified by a counterexample
  - A theorem imported via an interpretation
  - A theorem shown by a metatheorem

Transformers

- A **transformer** is a function that maps the expressions of a language L to the expressions of a language L'
  - Usually, $L \leq L'$, $L' \leq L$, or $L = L'$

- A transformer can be used to represent an expression transforming operation such as an evaluator, a simplifier, a rewrite rule, a rule of inference, a decision procedure, or an interpretation of one language in another

- Sound transformers can be:
  - Generated from theorems (e.g., theorem macetes)
  - Constructed from other transformers using certain constructors (e.g., compound macetes)
  - Obtained by instantiating abstract transformers (e.g., algebraic and order processors)
  - Manually defined and verified

Interpretations

- Theory interpretations can be used to:
  - Transport theorems, definitions, and profiles
  - Instantiate theories
  - Compare the strength of theories
  - Show relative consistency of theories
  - Show theory extension conservativity

- Logic interpretations can be used to interpret a theory in one logic in a theory of another logic