Automated reasoning 2018

Lecture 1: Introduction and background

Instructor: Ashutosh Gupta

IITB, India

Compile date: 2018-07-17

Topic 1.1

What is automated reasoning?

Automated reasoning (logic)

We will use reasoning
and logic synonymously.

- Have you ever said to someone "be reasonable"?
  - whatever your intuition was that is reasoning

- Why we care?

  Logic is the **calculus** of computer science

Example: applying logic

Logic is about inferring conclusions from given premises

Example 1.1

1. *Humans are mortal*
2. *Socrates is a human*

Socrates is mortal ✓

1. *Apostles are twelve*
2. *Peter is an apostle*

Peter is twelve ✗

Invalid reasoning?

Automated reasoning aims to

enable machines to

identify the valid reasoning!!

Automated reasoning is a backbone technology!!

Applications in verification, synthesis, solving NP-hard problems, and so on.

Automated reasoning for verification tools are like engines for the cars.

Topic 1.2

Spectra of logic

Propositional logic (PL)

Propositional logic

- deals with propositions,
- only infers from the structure over the propositions, and
- does not look inside the propositions.

Example 1.2

Is the following argument valid?
If the seed catalog is correct then if seeds are planted in April then the flowers bloom in July. The flowers do not bloom in July. Therefore, if seeds are planted in April then the seed catalog is not correct.

Let us symbolize our problem
If c then if s then f. not f. Therefore, if s then not c.

- c = *the seed catalogue is correct*
- s = *seeds are planted in April*
- f = *the flowers bloom in July*

PL reasons over propositional symbols and logical connectives

First-order logic (FOL)

First-order logic

- looks inside the propositions,
- much more expressive,
- includes parameterized propositions and quantifiers over individuals, and
- can express lots of interesting math.

Example 1.3

Is the following argument valid?
Humans are mortal. Socrates is a human. Therefore, Socrates is mortal.

In the symbolic form,
For all x if $H(x)$ then $M(x)$. $H(s)$. Therefore, $M(s)$.

- $H(x) = x$ *is a human*
- $M(x) = x$ *is mortal*
- $s =$ *Socrates*

FOL is not the most general logic.
Many arguments can not be expressed in FOL

Logical theories

In a theory, we study validity of FOL arguments under some specialized assumptions (called axioms).

Example 1.4

The number theory uses symbols $0, 1, .., <, +, \cdot$ *with specialized meanings*

The following sentence has no sense until we assign the meanings to $>$ *and* $\cdot$

$$\forall x\, \exists p\, (p > x \wedge (\forall v_1\, (v_1 > 1 \Rightarrow \forall v_2\; p \neq v_1 \cdot v_2)))$$

Under the meanings *it says that there are arbitrarily large prime numbers.*

In the earlier example, we had no interpretation of predicate 'x is human'. Here we precisely know what is predicate '$x < y$'.

Commentary: The specialized meaning are defined using axioms. For example, the sentence $\forall x.\ 0 + x = x$ describes one of the properties of 0 and $+$.

Higher-order logic (HOL)

Higher-order logic

- includes quantifiers over "anything",
- consists of hierarchy first order, second order, third order and so on,
- most expressive logic.

Example 1.5

$$\forall P \forall x.\ (P(x) \vee \neg P(x))$$

The quantifier is over proposition P. Therefore, the formula belongs to the second-order logic.

Commentary: The first quantifier is not allowed in the first-order logic. the first-order logic quantifies over individuals, the second-order logic quantifies over sets, the third-order logic quantifies over set of sets, and so on.

Topic 1.3

Satisfiability problem

Example: Satisfiability problem

Let x, y be rational variables.

Choose a value of x and y such that the following formula holds true.

$$x + y = 3$$

We say

$$\underbrace{\{x \mapsto 1, y \mapsto 2\}}_{\text{model}} \models \underbrace{x + y = 3}_{\text{formula}}.$$

Commentary: We are not calling x and y rational numbers. They are not numbers. They are symbols that can hold numbers.

Reasoning == Satisfiability problem

All reasoning problems can be reduced to satisfiability problems.

Often abbreviated to **SAT problem**

Exercise 1.1

How to convert checking a valid argument into a satisfiability problem?

Example: SAT problem(contd.)

Let x, y be rational variables.

Choose a value of x and y such that the following formula holds true.

$x + y = 3 \wedge y > 10 \wedge x > 0$ theory formulas Easy

$x + y = 3 \wedge y > 10 \wedge (x > 0 \vee x < -4)$ quantifier-free formulas Hard

$\forall y.\ x + y = 3 \wedge y > 10 \wedge (x > 0 \vee x < -4)$ quantified formulas Very Hard

Commentary: The above are increasingly hard classes of satisfiability problems. The names used for hardness are informal to minimize jargon.

Quantifier-free formulas

Quantifier-free formulas consists of

- theory atoms
- Boolean structure

Example 1.6

$$x + y = 3 \wedge y > 10 \wedge (x > 0 \vee x < -4)$$

Theory atoms

Boolean Structure

Commentary: We typically assume $\wedge$(conjunction) occurs in any formula. We say a formula has Boolean structure if it has $\vee$(disjunction).

Propositional formulas

Propositional formulas are a special case, where the theory atoms are Boolean variables.

Example 1.7

Let p_1, p_2, p_3 be Boolean variables

$$p_1 \wedge \neg p_2 \wedge (p_3 \vee p_2)$$

A satisfying model:

$$\{p_1 \mapsto 1, p_2 \mapsto 0, p_3 \mapsto 1\} \models p_1 \wedge \neg p_2 \wedge (p_3 \vee p_2)$$

A bit of jargon

- Solvers for quantifier-free propositional formulas are called

SAT solvers.

- Solvers for quantifier-free formulas with the other theories are called

SMT solvers.

SMT = satisfiability modulo theory

Theory solvers

SMT solvers are divided into two components.

- SAT solver: it solves the Boolean structure
- Theory solver: it solves the theory constraints

Example 1.8

Let x,y be rational variables.

$$x + y = 3 \wedge y > 10 \wedge x > 0$$

Since the formula has no $\vee$ (disjunction), a solver of linear rational arithmetic can find satisfiable model using simplex algorithm.

Commentary: Simplex can not directly handle the formulas that have disjunctions.

Quantified formulas

Quantified formulas also include quantifiers.

Example 1.9

The following formulas says: give x such that for each y the body holds true.

$$\forall y.\ \underbrace{(x + y = 3 \Rightarrow y > 10 \wedge (x > 0 \vee x < -4))}_{Body}$$

A satisfying model:

$$\{x \to 1\} \models \forall y.\ (x + y = 3 \Rightarrow y > 10 \wedge (x > 0 \vee x < -4))$$

Commentary: Since y is quantified, the model does not assign value to y.

Exercise

Exercise 1.2

Give satisfying assignments to the following formulas

- $\neg p_1 \wedge (p_1 \vee \neg p_2)$
- $x < 3 \wedge y < 1 \wedge (x + y > 5 \vee x - y < 3)$
- $\forall x\ (x > y \Rightarrow \exists z\ (2z = x))$

Topic 1.4

Course contents and logistics

Content of the course

We will study the following topics

- Background: first-order logic (FOL) basics
- SAT solvers: satisfiability solvers for propositional logic
- SMT solvers: satisfiability modulo theory solvers
- Decision procedures: algorithms for solving theory constraints
- Solvers for quantifiers
  - SMT+quantifier
  - Saturation solvers: FOL solvers
- Interactive theorem prover for higher order logics

Evaluation and website

- Programming assignments: 40%
- Quizzes : 10% (5% each)
- Midterm : 15% (1 hour)
- Presentation: 10% (15 min) [topics will be floated in the class]
- Final: 25% (2 hour)

For the further information

`http://www.cse.iitb.ac.in/~akg/courses/2018-automated-reasoning`

All the assignments and slides will be posted on the website.

Please read the conditions to attend the course. They are on the website.

Topic 1.5

First-order Logic

First-order logic(FOL)

$$\text{First-order logic(FOL)} \\ = \\ \text{propositional logic} + \text{quantifiers over individuals} + \text{functions/predicates}$$

"First" comes from this property

Example 1.10

Consider the following argument:
Humans are mortal. Socrates is a human. Therefore, Socrates is mortal.

In symbolic form,
$\forall x.(H(x) \Rightarrow M(x)) \wedge H(s) \Rightarrow M(s)$

- $H(x) = x$ *is a human*
- $M(x) = x$ *is mortal*
- $s =$ *Socrates*

A note on FOL syntax

The FOL syntax may appear non-intuitive and cumbersome.

FOL requires getting used to it like many other concepts such as complex numbers.

Connectives and variables

An FOL consists of three disjoint kinds of symbols

- variables
- logical connectives
- non-logical symbols : function and predicate symbols

Variables

We assume that there is a set **Vars** of countably many variables.

- Since **Vars** is countable, we assume that variables are indexed.

$$\textbf{Vars} = \{x_1, x_2, \ldots,\}$$

- The variables are just names/symbols without any inherent meaning
- We may also sometimes use x, y, z to denote the variables

Now forget all the definitions of the propositional logic. We will redefine everything and the new definitions will subsume the PL definitions.

Logical connectives

The following are a finite set of symbols that are called logical connectives.

formal name	symbol	read as	
true	$\top$	top	0-ary
false	$\bot$	bot	
negation	$\neg$	not	unary
conjunction	$\wedge$	and	binary
disjunction	$\vee$	or	
implication	$\Rightarrow$	implies	
exclusive or	$\oplus$	xor	
equivalence	$\Leftrightarrow$	iff	
equality	$\approx$	equals	binary predicate
existential quantifier	$\exists$	there is	quantifiers
universal quantifier	$\forall$	for each	
open parenthesis	(		punctuation
close parenthesis	)		
comma	,		

Non-logical symbols

FOL is a parameterized logic

The parameter is a signature $\mathbf{S} = (\mathbf{F}, \mathbf{R})$, where

- **F** is a set of function symbols and
- **R** is a set of predicate symbols.

Each symbol has arity ≥ 0

F and **R** may either be finite or infinite.

Each **S** defines an FOL.
We say, consider an FOL with signature $\mathbf{S} = (\mathbf{F}, \mathbf{R})$ ...

We write $f/n \in \mathbf{F}$ and $P/k \in \mathbf{R}$ to explicitly state the arity

With $n = 0$, f is called constant
With $k = 0$, P is called propositional variable

Syntax : terms

Definition 1.1

For signature $\mathbf{S} = (\mathbf{F}, \mathbf{R})$, $\mathbf{S}$-terms $T_{\mathbf{S}}$ are given by the following grammar:

$$t \triangleq x \mid f(\underbrace{t, \dots, t}_{n}),$$

where $x \in \mathbf{Vars}$ and $f/n \in \mathbf{F}$.

Example 1.11

Consider $\mathbf{F} = \{c/0, f/1, g/2\}$.
The following are terms

- $f(x_1)$
- $g(f(c), g(x_2, x_1))$
- c
- x_1

You may be noticing some similarities between variables and constants

Some notation:

- Let $\vec{t} \triangleq t_1, .., t_n$

Infix notation

We may write some functions and predicates in infix notation.

Example 1.12

we may write $+(a, b)$ *as* $a + b$ *and similarly* $< (a, b)$ *as* $a < b$.

Compact notation for terms

Since we know arity of each symbol, we need not write "," "(", and ")" to write a term unambiguously.

Example 1.13

$f(g(a,b),h(x),c)$ can be written as fgabhxc.

Exercise 1.3

*Consider $\mathbf{F} = \{f/3, g/2, h/1, c/0\}$ and $x, y \in$ **Vars**.*
Insert parentheses at appropriate places in the following terms.

- $hc =$
- $gxc =$
- $fhxhyhc =$
- $fx =$

Exercise 1.4

Give an algorithm to insert the parentheses

Commentary: We will not use the compact notation in the course. It makes the formulas very difficult to read.

Syntax: atoms

Definition 1.2

S-*atoms* $A_{\mathbf{S}}$ *are given by the following grammar:*

$$a \triangleq P(\underbrace{t, \dots, t}_{n}) \mid t \approx t \mid \bot \mid \top,$$

where $P/n \in \mathbf{R}$.

Exercise 1.5

Consider $\mathbf{F} = \{s/0\}$ *and* $\mathbf{R} = \{H/1, M/1\}$
Is the following an atom?

- $H(x)$
- s
- $M(s)$
- $H(M(s))$

Syntax: formulas

Definition 1.3

S-*formulas* $\mathbf{P_S}$ *are given by the following grammar:*

$$F \triangleq a \mid \neg F \mid (F \wedge F) \mid (F \vee F) \mid (F \Rightarrow F) \mid (F \Leftrightarrow F) \mid (F \oplus F) \mid \forall x.(F) \mid \exists x.(F)$$

where $x \in \mathbf{Vars}$.

Example 1.14

Consider $\mathbf{F} = \{s/0\}$ *and* $\mathbf{R} = \{H/1, M/1\}$

The following is a $(\mathbf{F}, \mathbf{R})$-*formula:*

$$\forall x.(H(x) \Rightarrow M(x)) \wedge H(s) \Rightarrow M(s)$$

FOL formulas

Exercise 1.6

Convert the following english sentences into $(\{zero/0, succ/1\}, \{\})$*-formulas.*

- *There is always a successor of a natural number.*
- *If any two numbers have same successor, then the two numbers are equal*
- *Zero is not successor of any number.*

Topic 1.6

FOL terminology

Subterm and subformulas

Definition 1.4

A term t is subterm of term t', if t is a substring of t'.

Definition 1.5

A formula F is subformula of formula F', if F is a substring of F'.

Some terminology

We may not mention **S** if from the context the signature is clear.

Closed terms and quantifier free

Definition 1.6

A closed term is a term without variable. Let $\hat{T}_{\mathbf{S}}$ be the set of closed $\mathbf{S}$-terms. Sometimes closed terms are also referred as ground terms.

Definition 1.7

A formula F is quantifier free if there is no quantifier in F.

Free variables

Definition 1.8

A variable $x \in \mathbf{Vars}$ is free in formula F if

- $F \in A_{\mathbf{S}}$: *x occurs in* F,
- $F = \neg G$: *x is free in* G,
- $F = G \circ H$: *x is free in G or H, for some binary operator* $\circ$, *and*
- $F = \exists y.G$ *or* $F = \forall y.G$: *x is free in G and* $x \neq y$.

Let $FV(F)$ *denote the set of free variables in* F.

Exercise 1.7

Is x free?

- $H(x)$
- $H(y)$
- $\forall x.H(x)$
- $x \approx y \Rightarrow \exists x.G(x)$

Sentence

Definition 1.9

*A variable $x \in$ **Vars** is bounded in formula F if x occurs in F and x is not free. In $\forall x.G$ ($\exists x.G$), we say the quantifier $\forall x$ ($\exists x$)has scope G and bounds x.*

Definition 1.10

A formula F is a sentence if it has no free variable.

Exercise 1.8

Is the following formula a sentence?

- $H(x)$
- $\forall x.H(x)$
- $x \approx y \Rightarrow \exists x.G(x)$
- $\forall x.\exists y.\ x \approx y \Rightarrow \exists x.G(x)$

Topic 1.7

Problems

FOL formulas

Exercise 1.9

Convert the following english sentences into FOL-formulas. Choose signature appropriately.

- *There is a cap for every pen.*
- *There is someone such that if the one drinks, then everyone drinks.*
- *There is an irrational number such that power of it to an irrational exponent is rational.*

FOL Syntax

Exercise 1.10

Which of the following sentences are well-formed FOL **S***-formulas, where* $\mathbf{S} = (\{c/0, f/1, g/2\}, \{P/0, Q/1, R/2\})$*?*

- c
- P
- $R(x, y) \Rightarrow Q(f(x, z))$
- $\forall x.\ Q(y) \land P$
- $\forall x.\ f(c, x) \approx c$
- $f(c) \approx P$

End of Lecture 1