Automatic Feedback on Logic Problems about Real Numbers and Integers

Antti Valmari
University of Jyväskylä, Faculty of Information Technology

1 Introduction

Goal: immediate advanced automatic feedback on early university math problems

- more feedback than just right / wrong
  - counter-examples
  - problem-specific format requirements, e.g., `f_polynomial`

$n(2n-1)(2n+1) = n(2n^2-1)$

Relation does not hold when $n = 1$

left $= 3$
right $= 1$

$n(2n-1)(2n+1) = n(4n^2-1)$

That is correct!

$n(2n-1)(2n+1) = n(4n^2-1)$

The final expression must be in polynomial form. The problem is: $n(4n^2-1)$

$n(2n-1)(2n+1) = n(4n^2-1) = 4n^3-n$

That is correct, and in polynomial form!

⇒ students can make more progress on their own

- mainstream math education tool research seems to not know anything similar [Kinnear & al. 2024]

(Part of) our earlier work

Click either button to see the feedback by MathCheck.

```
log^2(a+b) = (log(a+b))^2 = (log a+b)^2
```

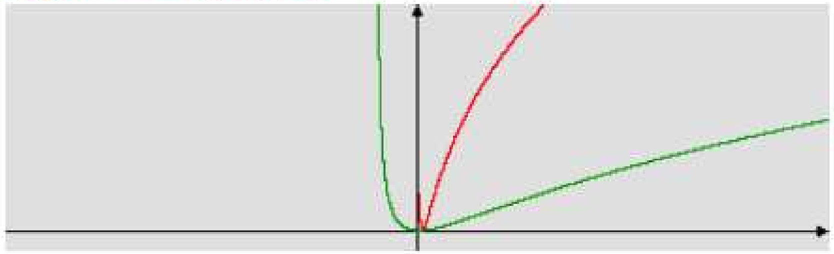
new tab or to the right 1

Typing instructions: arithmetic comparisons basic logic

$\log^2(a+b) = (\log(a+b))^2 = (\log a + b)^2$

Relation does not hold when $a = 0$ and $b = 1$

left $=$ 0
right $=$ undefined



$-10 \leq a \leq 10$ and $-0.2 \leq \uparrow \leq 2.2$

- try many combinations of values of variables
  - fundamentally unreliable
  - in practice reliable enough for $= \leq$ etc. -chains, excluding special cases (e.g., individual zero divisors)
- exceptionally math-like syntax
  - e.g., $\sin 2x = 2 \sin x \cos x$

This paper: three logic modes

- real numbers
- (natural numbers)
- integers

Restriction to "almost linear" formulas

- can be reliably automated with reasonable running times
- linear expression:
  - e.g., $12z - 5y + x + 20$
  - no xy, x^2, ...
- somewhat more general cases can be transformed to formulas with linear expressions
  - $\frac{x+1}{x-1} = 2 \quad \Leftrightarrow \quad x \neq 1 \wedge x + 1 = 2(x-1)$
  - $3|x| \geq |x-3| + 5 \quad \Leftrightarrow$
    $x < 0 \wedge -3x \geq -(x-3) + 5 \vee 0 \leq x < 3 \wedge 3x \geq -(x-3) + 5 \vee x \geq 3 \wedge 3x \geq (x-3) + 5$
- a big restriction — yet allows a lot

To properly deal with undefined expressions, we use the logic in [Valmari & Hella 2023]

- in this paper: division by zero
- elsewhere in MathCheck: indexing an array out of bounds, ...

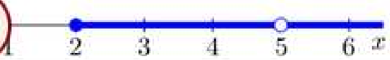
2 “Almost Linear” Logic Problems on Real Numbers

Line segment problems

Write a short formula so that the formula is true if and only if x is in the blue area.

Hint

Say within your answer that x is not 5.




```
2 <= x < 5 \/ x > 5
```

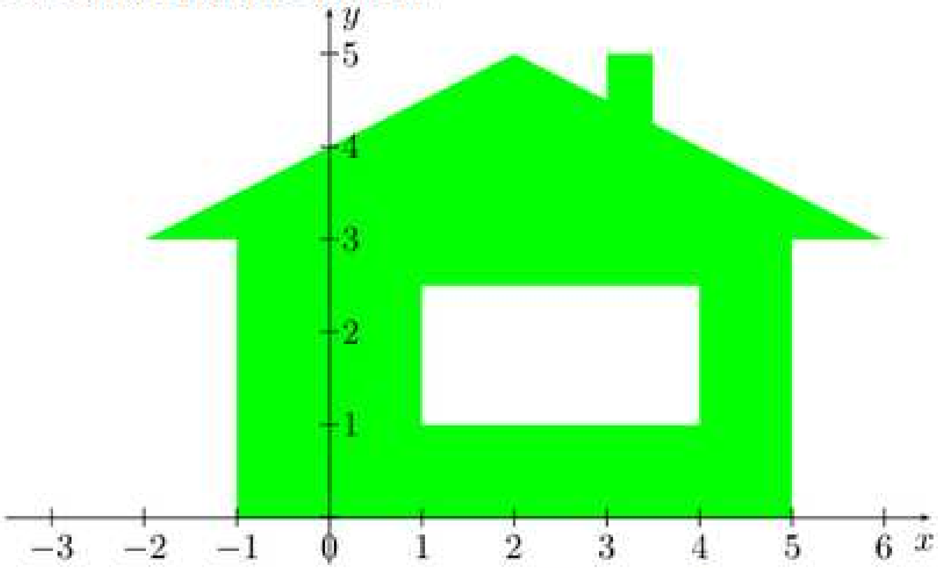
new tab or to the right 4

model-answer ⇔ $2 \leq x < 5 \vee x > 5$

The complexity of the final expression is 9, while it must be at most 7.

The cottage problem

Write a formula that represents the green area. All points on the borderlines between white and green are green. You may first design and check each part of the cottage in the answer boxes that are further below.




```
0 <= 2y <= x + 8 /\ (2y <= 12-x \/ 3 <= x <= 7/2 /\ y <= 5 ) /\ (-1 <= x <= 5 \/ y >= 3)
/\ !(1 < x < 4 /\ 1 < y < 2 1/2)
```

Challenging system of (in)equations

$\Rightarrow$ a small language for expressing reasoning

Solve the inequation $3|x| \geq |x-3|+5$. Start with ⇒, ⇐, ⇔, ≡, or `subproof`.

```
subproof x < 0 /\ -3x >= -x+3 + 5 <=> x < 0 /\ x <= -4 <=> x <= -4 subend
subproof assume 0 <= x < 3; 3x >= -x+3 + 5 <=> x >= 2 <=> 2 <= x < 3 subend
subproof 3x >= x-3 + 5 <=> x >= 1 subend
original <=> x <= -4 \/ 2 <= x < 3 \/ x >= 3 /\ x >= 1 <=> x <= -4 \/ x >= 2
```

new tab or to the right 11

$3|x| \geq |x-3|+5$

Subproof

$x < 0 \land -3x \geq -x+3+5 \Leftrightarrow x < 0 \land x \leq -4 \Leftrightarrow x \leq -4$

Subend

Subproof

Assume $0 \leq x < 3$

$3x \geq -x+3+5 \Leftrightarrow x \geq 2 \Leftrightarrow 2 \leq x < 3$

Subend

Subproof

$3x \geq x-3+5 \Leftrightarrow x \geq 1$

Subend

Original $\Leftrightarrow x \leq -4 \lor 2 \leq x < 3 \lor x \geq 3 \land x \geq 1 \Leftrightarrow$
$x \leq -4 \lor x \geq 2$

3 “Almost Linear” Logic Problems on Integers

Motivation

- sometimes the numbers are definitely integers, not real numbers
  - what is known on the value of i in the beginning of line 3?

```
1  void SelectionSort( array_type & A ){
2    for( unsigned i = 0; i+1 < A.size(); ++i ){
3      unsigned p = i;
4      for( unsigned j = i+1; j < A.size(); ++j ){
5        if( A[j].x < A[p].x ){ p = j; }
6      }
7      elem_type tmp = A[i]; A[i] = A[p]; A[p] = tmp;
8    }
9  }
```

  - both $0 \le i < n-1$ and $0 \le i \le n-2$ are correct

⇒ problems highlighting the differences of integers and reals are useful
  - the volleyball problem

- meaningful quantifier problems
  - without using div and mod, write a formula that says that n is divisible by 7

    $\exists k : n = 7k$

Counter-example production can sometimes be harnessed to reveal some number

- want to know a number that can be represented as a sum of two cubes of natural numbers in two different ways?

⇒ claim MathCheck that there is no such number!

- the teacher has secretly defined $C(n,m)$ as $n = m = 0 \vee \ldots \vee n = 16 \wedge m = 4096$
- the student should first write that x is the sum of the cubes of n and m
  - e.g., $\exists c : C(n,c) \wedge C(m,x-c)$
  - this formula becomes $S(n,m,x)$
- the student should next write that x cannot be represented as both $n^3 + m^3$ and $k^3 + h^3$, where these are different
  - e.g., $\neg(S(n,m,x) \wedge S(k,h,x) \wedge n \neq h \neq m)$

Prove that if x does not occur free in P, then $\exists x : P \wedge Q(x) \Leftrightarrow P \wedge \exists x : Q(x)$

- simplify both sides when $P \equiv$ **F**
- simplify both sides when $P \equiv$ **T**

Provide a counter-example to $\forall x : (P(x) \vee Q(x)) \Leftrightarrow (\forall x : P(x)) \vee (\forall x : Q(x))$

4 DFA Implementat. of Integer Presburger Arithmetic

Can be used for illustrating many computer science topics

An *unbounded* integer as a bit string

$$83 = \boxed{0\,|\,1\,|\,1\,|\,0\,|\,0\,|\,1\,|\,0\,|\,1\,|\,0\,|\,0\,|\,0\,|\,0\,|\,0}$$

sign 2^0 2^1 2^4 2^6 any number of 0's

- the empty string represents 0
- the string 1 represents -1
- if the sign is 1, each further 1 *subtracts* the weight
  - e.g., the string 11 represents -2

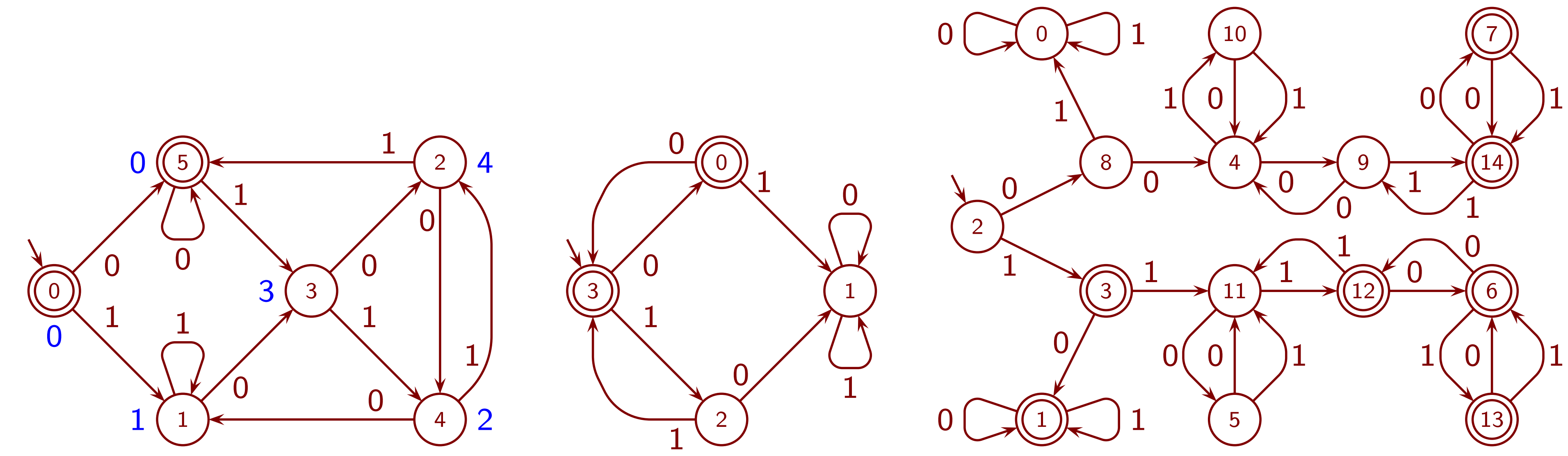
$\Rightarrow$ extending a string with 0 never changes its value

A triple of unbounded integers as a bit string

$$(2025, -9, -10) = \boxed{0\,1\,1\,|\,1\,0\,1\,|\,0\,0\,0\,|\,0\,0\,0\,|\,1\,1\,1\,|\,0\,0\,0\,|\,1\,0\,0\,|\,1\,0\,0\,|\,1\,0\,0\,|\,1\,0\,0\,|\,1\,0\,0\,|\,1}$$

A formula is represented as

- an ordered list of (at least) its free variables
- the set of satisfying assignments as a set of bit strings as a DFA



$x \bmod 5 = 0$ $\qquad$ $x = y$ $\qquad$ $x < y$

Operations

$\neg$	trivial
$\wedge\ \vee\ \rightarrow\ \leftrightarrow$	product automata — advanced data structures and algorithms
$\exists x\ \forall x$	subset construction — very challenging algorithms!
$+$	a 33-state DFA $\qquad -37x$ $\quad$ a 119-state DFA $\qquad \ldots$
$\cdot$	impossible, due to Gödel's incompleteness theorem

5 Concluding Remarks

All well-organized pedagogical experiments are old

- a re-organization of the university stopped my research on teaching tools

$\Rightarrow$ moved to a more sympathetic university

- I and a math teacher use the tool a lot, but lack the skills for making pedagogical experiments

Student feedback

- until Covid-19: some very positive, some positive, next to none negative
- since Covid-19: next to none feedback

Thank You for attention! **Questions, discussion?**