STACK developer update, EAMS 2021

Chris Sangwin

School of Mathematics
University of Edinburgh

June 2021

2020-21

Unprecedented interest in teaching online....

2020-21

Unprecedented interest in teaching online....
Focus in summer 2020

- STACK user workshops
- Assessment of Proof (with R. Bickerton)

2020-21

Unprecedented interest in teaching online....
Focus in summer 2020

- STACK user workshops
- Assessment of Proof (with R. Bickerton)

(Not a lot of new features in summer of 2020!)

STACK 4.3.8: December 2020

- Introduce "context variables"
- Make test case construction easier.
- Internal reorganisation of answer tests.
- Add in house styles
  - Proof
  - HELM

Version 4.3.9: summer 2021

- Student's input now allows coordinates (x, y).
- Add in warnings of language mismatch.
- Add in warnings where the answer test needs a raw input.
- Expand `rand` to now accept sets.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.
- Make subtle distinctions.
  e.g. accept answer of $0 - i$ but reject $1/i$.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.
- Make subtle distinctions.
  e.g. accept answer of $0 - i$ but reject $1/i$.
  Select and use some algebraic rules but not others.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.
- Make subtle distinctions.
  e.g. accept answer of $0 - i$ but reject $1/i$.
  Select and use some algebraic rules but not others.
- Making subtle distinctions is necessary for helpful feedback.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.
- Make subtle distinctions.
  e.g. accept answer of $0 - i$ but reject $1/i$.
  Select and use some algebraic rules but not others.
- Making subtle distinctions is necessary for helpful feedback.

I will assume familiarity with writing STACK questions.
Apologies to new STACK users.

Roll your own answer test

- Solving problems at the mathematical level is a key goal.
- Make subtle distinctions.
  e.g. accept answer of $0 - i$ but reject $1/i$.
  Select and use some algebraic rules but not others.
- Making subtle distinctions is necessary for helpful feedback.

I will assume familiarity with writing STACK questions.
Apologies to new STACK users.
(New users: sorry if the answer test is technical!
Algebraic equivalence is fine 90% of the time...)

Basic CAA process

1. The teacher poses a question.

Basic CAA process

1. The teacher poses a question.
2. The student answers.

Basic CAA process

1. The teacher poses a question.
2. The student answers.
3. Establish properties.

Basic CAA process

1. The teacher poses a question.
2. The student answers.
3. Establish properties.
4. Generate *outcomes*, e.g. mark/feedback/stats.

Tests: 1. Syntactic tests.

The *form* of the answer.

Tests: 1. Syntactic tests.

The *form* of the answer.
Factor command returns $(x-2)^2$.

Tests: 1. Syntactic tests.

The *form* of the answer.
Factor command returns $(x-2)^2$.

$$(x-2)(x-2), \quad (x-2)^2, \quad (2-x)^2, \quad 4\left(1-\frac{x}{2}\right)^2.$$

Tests: 1. Syntactic tests.

The *form* of the answer.
Factor command returns $(x-2)^2$.

$$(x-2)(x-2), \quad (x-2)^2, \quad (2-x)^2, \quad 4\left(1-\frac{x}{2}\right)^2.$$

Recognising a “factored" expression:

1. A product
2. of powers
3. of distinct
4. irreducible terms.

(“Irreducible?”)

Tests: 2. Equivalence tests

Characterized by a *family* of answers which pass.

1. Exactly the same. (String match)

Tests: 2. Equivalence tests

Characterized by a *family* of answers which pass.

1. Exactly the same. (String match)
2. Algebraic equivalence.

Tests: 2. Equivalence tests

Characterized by a *family* of answers which pass.

1. Exactly the same. (String match)
2. Algebraic equivalence.
3. Substitution equivalence.

$$a^2 + b^2 = c^2 \quad \equiv \quad x^2 + y^2 = z^2.$$

Tests: 2. Equivalence tests

Characterized by a *family* of answers which pass.

1. Exactly the same. (String match)
2. Algebraic equivalence.
3. Substitution equivalence.

$$a^2 + b^2 = c^2 \quad \equiv \quad x^2 + y^2 = z^2.$$

4. Same type.

A surprisingly useful meaning of “same"

We want

$$2x + 1 = 1 + 2x$$

BUT

$$x + x + 1 \neq 1 + 2x.$$

A surprisingly useful meaning of “same"

We want

$$2x + 1 = 1 + 2x$$

BUT

$$x + x + 1 \neq 1 + 2x.$$

Equivalence up to associativity and commutativity of $+$ and $\times$.

Current limitations

Calculate the following power and write your answer in the Cartesian form: $\left(-\frac{i-1}{\sqrt{2}}\right)^{-14}$.

Current limitations

Calculate the following power and write your answer in the Cartesian form: $\left(-\frac{i-1}{\sqrt{2}}\right)^{-14}$.

The power is randomly generated.
The correct answer is $-i$.

Students' responses

Year 1, "Proofs and Problem Solving" course.

Frequency	Response
50 (72.46%)	`-i`
4 (5.80%)	`0-i`
3 (4.35%)	`0-1*i`
2 (2.90%)	`1/i`
1 (1.45%)	`cos(5/2*pi)-sin(5/2*pi)*i`
1 (1.45%)	`0+1*i`
1 (1.45%)	`i`
1 (1.45%)	`cos(-5/2*pi)+i*sin(-5/2*pi)`
1 (1.45%)	`-1/32`
1 (1.45%)	`1`
1 (1.45%)	`-1*i`

The problem

Inconsistency

- "Cartesian form" is $x + i\,y$.
- Do not write $0+?$ and $1\times?$.

The problem

Inconsistency

- "Cartesian form" is $x + i\,y$.
- Do not write $0+?$ and $1\times?$.

About 10% of the cohort typed in something like `0-1*i` .

The problem

Inconsistency

- "Cartesian form" is $x + i\,y$.
- Do not write $0+?$ and $1\times?$.

About 10% of the cohort typed in something like `0-1*i` .

What do you want to accept?

The problem

Inconsistency

- "Cartesian form" is $x + i\,y$.
- Do not write $0+?$ and $1\times?$.

About 10% of the cohort typed in something like `0-1*i` .

What do you want to accept?
But, don't accept anything equivalent, e.g. `1/i`.

Problem 2

Simplify $\frac{4x^2-4y^2}{-6y-6x}$.

The correct answer is $-\frac{2}{3}(x-y)$.

Problem 2

Simplify $\frac{4x^2-4y^2}{-6y-6x}$.

The correct answer is $-\frac{2}{3}(x-y)$.
Up to commutativity and associativity we have
$-(2/3)(x-y)$, $(-2/3)(x-y)$, $(-2(x-y))/3$, $-(2(x-y))/3$,
$-2(x-y)/3$, and $2(-(x-y)/3)$.

Two problematic groups

The second class includes $(2/-3)(x-y)$, $(2(x-y))/(-3)$, and $2(x-y)/(-3)$.

Two problematic groups

The second class includes $(2/-3)(x-y)$, $(2(x-y))/(-3)$, and $2(x-y)/(-3)$.

The third class includes $(2(y-x))/3$, $(2/3)(y-x)$, and $2(y-x)/3$.

Solution

New answer test: AlgEquivNouns

Solution

New answer test: AlgEquivNouns

Problems with this approach

1. We can't keep introducing new answer tests!
2. Need much more flexible control.
   Question by question.

Key idea

Establish whether or not two parse trees belong to the same equivalence class.

Key idea

Establish whether or not two parse trees belong to the same equivalence class.

Solution: teachers define their own classes.

Rule-based answer test

1. Teacher chooses relevant rules.

Rule-based answer test

1. Teacher chooses relevant rules.
2. System repeatedly applies those rules across the tree until the expression stops changing.

Rule-based answer test

1. Teacher chooses relevant rules.
2. System repeatedly applies those rules across the tree until the expression stops changing.
3. The equivalence class is based on the final form.

Choose rules

Name	Rule
(ALG-TRANS)	
assAdd	Associativity of addition
assMul	Associativity of multiplication
comAdd	Commutativity of addition
comMul	Commutativity of multiplication
(ID-TRANS)	
zeroAdd	$0 + x \to x$
zeroMul	$0 \times x \to 0$
oneMul	$1 \times x \to x$
oneDiv	$\frac{x}{1} \to x$
onePow	$1^x \to 1$
idPow	$x^1 \to x$
zeroPow	$0^x \to 0$ if $x \neq 0$
zPow	$x^0 \to 1$ if $x \neq 0$

Name	Rule
(NEG-TRANS)	
negNeg	$-(-x) \to x$
negDiv	$y/(-x) \to -y/x$
(DIV-TRANS)	
recipMul	$x/a \times y/b \to (x\,y)/(a\,b)$
divDiv	$a/(b/c) \to a\,c/b$
(INT-ARITH)	
intAdd	Perform addition on integers
intMul	Perform multiplication on integers
intPow	Perform integer exponentiation

Results

Test	Frequency	Response
CASEqual	50 (72.46%)	`-i`
ID-TRANS	4 (5.80%)	`0-i`
	3 (4.35%)	`0-1*i`
	1 (1.45%)	`-1*i`
AlgEquiv	2 (2.90%)	`1/i`
	1 (1.45%)	`cos(5/2*pi)-sin(5/2*pi)*i`
	1 (1.45%)	`cos(-5/2*pi)+i*sin(-5/2*pi)`
(!Wrong!)	1 (1.45%)	`0+1*i`
	1 (1.45%)	`i`
	1 (1.45%)	`-1/32`
	1 (1.45%)	`1`

Notes

1. Need to ensure the algorithm terminates.
   Can't have "factor" and "expand" in the same set!

Notes

1. Need to ensure the algorithm terminates.
   Can't have "factor" and "expand" in the same set!
2. Rules don't have to be correct mathematics.

Conclusion

- Close to a release. Will be ready for 2021-22.

Conclusion

- Close to a release. Will be ready for 2021-22.
- Answer test: I did not know what I wanted when I started!

Conclusion

- Close to a release. Will be ready for 2021-22.
- Answer test: I did not know what I wanted when I started!
- I don't expect to know all use-cases.

Conclusion

- Close to a release. Will be ready for 2021-22.
- Answer test: I did not know what I wanted when I started!
- I don't expect to know all use-cases.

(New users: sorry if the answer test is technical!
Algebraic equivalence is fine 90% of the time...)