

Math Aware Search: Experimenting with a data driven math recognizer

June 9, 2022

George Labahn and Tia Ruza

University of Waterloo

Huawei-Waterloo Math Aware Search Project

Project has two components:

- Searching documents containing math expressions
 - described in previous 2 workshops
- Inputting mathematical expressions
 - described today

Which is easier?

$$\int \frac{(3x^2+2)\sin(x^3+2x-1)}{\cos(x^3+2x-1)^3 + 12\cos(x^3+2x-1)^2-3} dx$$

◆ Latex

```
\begin{equation}
\int \frac{(3x^2+2)\sin(x^3+2x-1)}{\cos(x^3+2x-1)^3 + 12\cos(x^3+2x-1)^2-3} dx
\end{equation}
```

◆ Maple

```
Int((3*x^2+2)*sin(x^3+2*x-1)/(cos(x^3+2*x-1)^3 + 12*cos(x^3+2*x-1)^2-3),x);
```

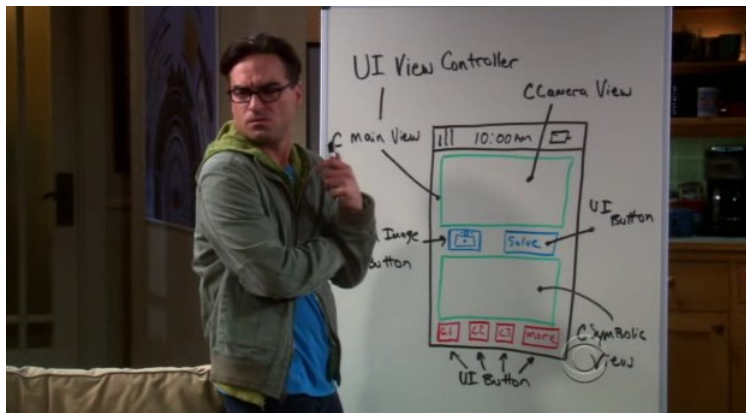
◆ Mathematica

```
Integrate[(3*x^2+2)*Sin[x^3+2*x-1]/(Cos[x^3+2*x-1]^3 + 12*Cos[x^3+2*x-1]^2-3),x]
```

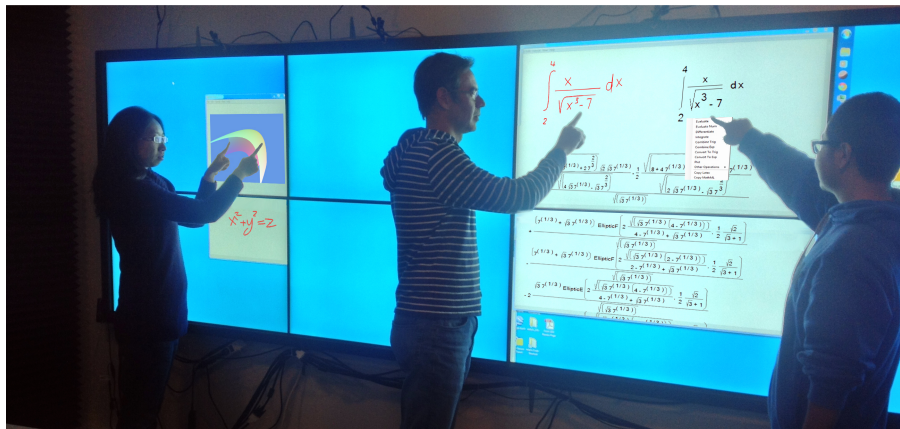
Doing Math on a Computer is challenging

- Entering mathematical concepts on a computer for **presentation, computation or search** is tedious and difficult
- Current interfaces of math input constrain the user to express their thoughts in a command line form
- Different systems use different languages
- Manipulation is not intuitive with keyboard and mouse

Interest from unusual sources e.g. entertainment



Interesting potential uses: e.g. Math on a Powerwall



Challenges I

- Text recognizers:
 - work with ASCII characters
 - assume input is one dimensional
 - depend on language specific dictionaries
- Math symbols more ambiguous than regular text, e.g.



1+1 or |t| ?



SyZ or 5x2 ?

Challenges II

- Math symbols also
 - have wide variation in size
 - do not necessarily follow baselinese.g. superscripts and subscripts
- Correct symbols recognition \nrightarrow unique math expressions.
e.g. is $u(x + y)$ u times $x + y$ or u applied to argument $x + y$?
- Rendering 2 dimensional math expressions with line-breaking.
- Users require interactivity: editing, gestures, sub-expressions

MathBrush Recognizer

- MathBrush recognizer (original)
 - Recognition is a “fuzzy” process
 - Return a ‘best bet’ along with alternatives
 - Easy to do corrections
 - Ability to see recognition as user writes
- Ability to edit expressions
- Ability for users to train their symbols
- Use context menus to do math

MathBrush: Math Recognizer

- Grammar based model
- Uses probabilistic decision reasoning
- Three step process : symbolic recognition, parsing and tree extraction
- Returns ranked alternatives for symbols and sub-expressions.
- Ability to train recognizer for symbols

CROHME Competition

- Competition on Recognition of Online Handwritten Mathematical Expressions (CROHME)
- Run in 2011, 2012, 2013, 2014, . . . , 2019
- MathBrush finished second in 2012. First place went to Vision Objects, a commercial outfit (myscript math recognizer)
- Three main recognition tasks in 2014
 - Isolated symbols
 - Full expressions
 - Matrices (new in 2014 – because of MathBrush)

Part II

Improving the Query-Driven Global Graph Attention Recognizer

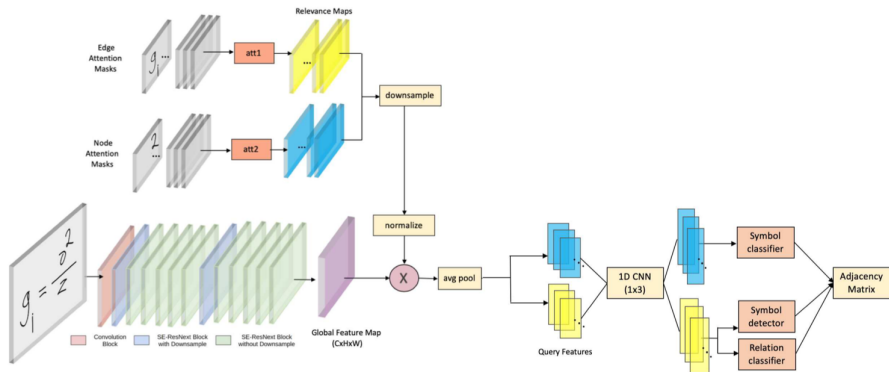
Why the QD-GGA Recognizer?

- Open-source
- Modular architecture, separating symbol recognition and structural parsing, which allows:
 - Alternative Structural Parses
 - Alternative Symbol Recognition
 - Potential matrices support
 - User-specific customization

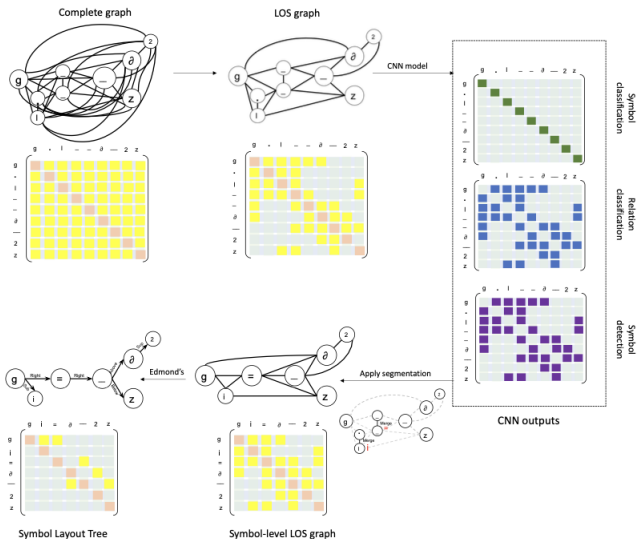
QD-GGA Recognizer: Overview

- Depends on a Convolutional Neural Network (CNN)
- Uses a Query-Driven Global Graph Attention (QD-GGA) model
- Adapted from Mathseer-extraction-pipeline of the Document and Pattern Recognition Lab (DPRL) from Rochester Institute of Technology

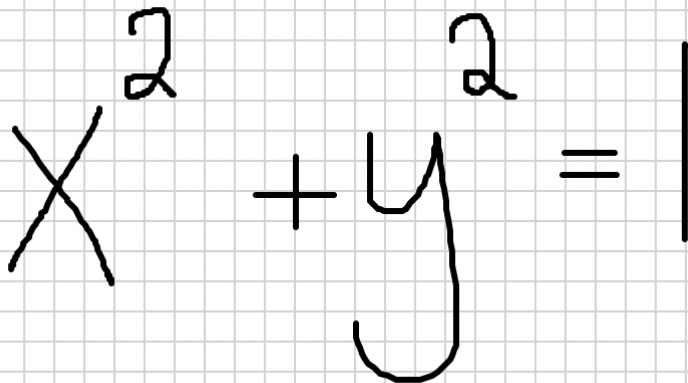
QD-GGA Recognizer: System Architecture



QD-GGA Recognizer: Parse Example



QD-GGA Recognizer: Parse Example



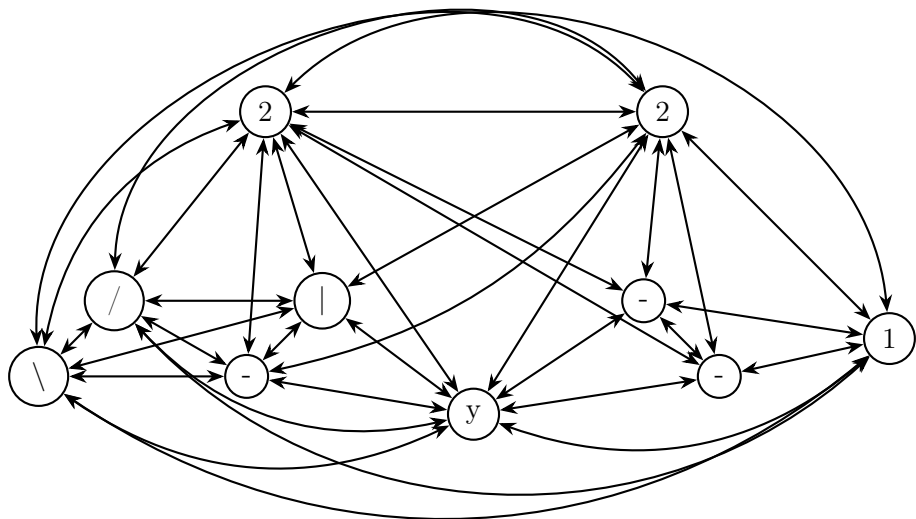
A handwritten equation $x^2 + y^2 = 1$ is displayed on a grid background. The equation is written in black ink. The 'x' is a simple cross, the 'y' is a tall vertical stroke with a curved bottom, and the '2' characters are small loops. The equals sign consists of two parallel horizontal lines, and the '1' is a single vertical line.

QD-GGA Recognizer: Model

The QD-GGA recognizer parses mathematical expressions as follows:

- 1 Create Line-of-Sight graph

QD-GGA Recognizer: Parse Example

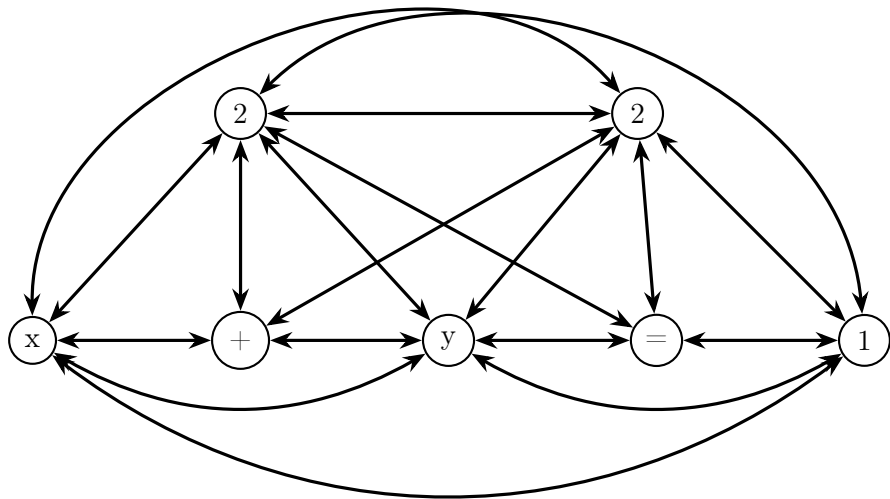


QD-GGA Recognizer: Model

The QD-GGA recognizer parses mathematical expressions as follows:

- ① Create Line-of-Sight graph
- ② Apply segmentation with the CNN

QD-GGA Recognizer: Parse Example



QD-GGA Recognizer: Model

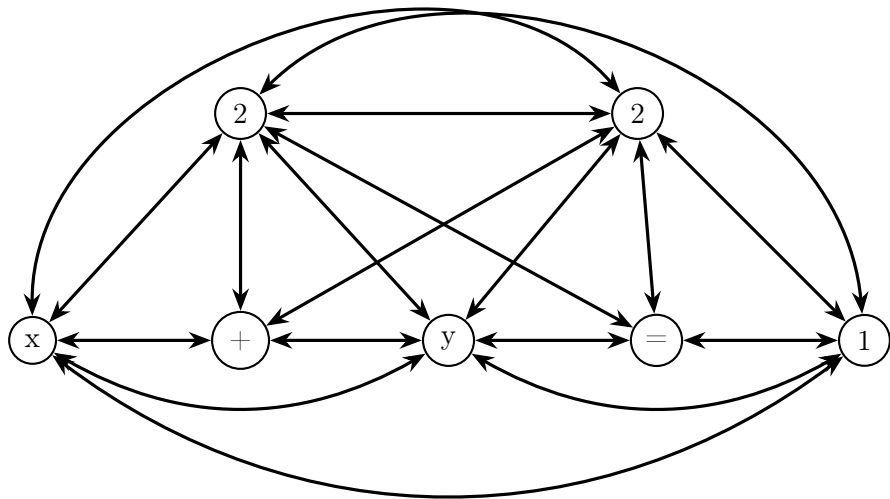
The QD-GGA recognizer parses mathematical expressions as follows:

- ① Create Line-of-Sight graph
- ② Apply segmentation with the CNN
- ③ Find symbol and relation probabilities with the CNN

Possible relations:

Right, Above, Below, Sup, Sub, Inside, No Relation

QD-GGA Recognizer: Parse Example



QD-GGA Recognizer: Model

The QD-GGA recognizer parses mathematical expressions as follows:

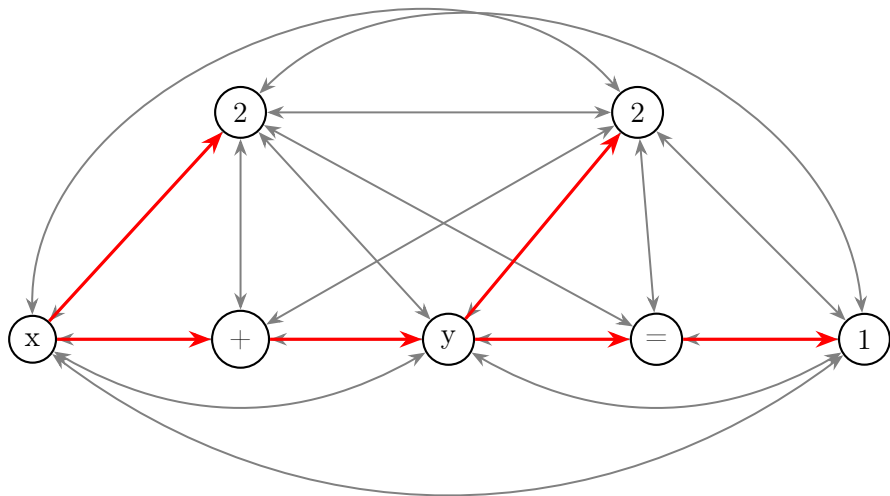
- ① Create Line-of-Sight graph
- ② Apply segmentation with the CNN
- ③ Find symbol and relation probabilities with the CNN

Possible relations:

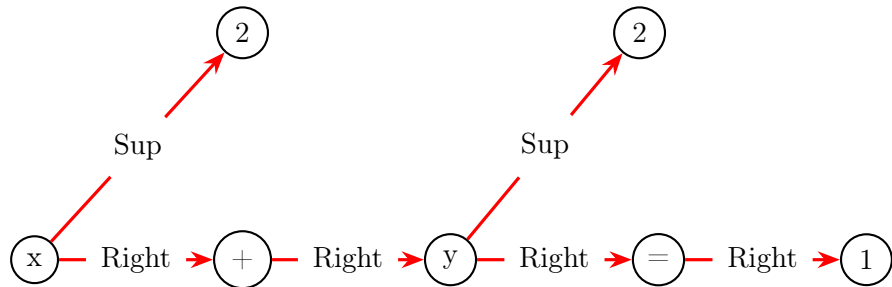
Right, Above, Below, Sup, Sub, Inside, No Relation

- ④ Determine maximum directed rooted spanning tree with Edmond's arborescence algorithm

QD-GGA Recognizer: Parse Example



QD-GGA Recognizer: Parse Example



QD-GGA Recognizer: Model

The QD-GGA recognizer parses mathematical expressions as follows:

- ① Create Line-of-Sight graph
- ② Apply segmentation with the CNN
- ③ Find symbol and relation probabilities with the CNN

Possible relations:

Right, Above, Below, Sup, Sub, Inside, No Relation

- ④ Determine maximum directed rooted spanning tree with Edmond's arborescence algorithm

Alternatives

BrushSearch Math Stack Exchange Close Gettable Sign Out

\int x^2 C^{-2} dx

Change Recognition

Alternatives..

- C
- 2
- e
- c
- z

Save and Close

ds comma separated

$$\int x^2 e^{-x^2} dx$$

QD-GGA Recognizer: Alternatives

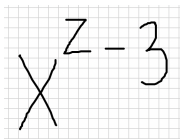
Progress has been made on a key step:

Correcting errors quickly through **Alternatives**

Given any list of nodes, recognizer gives up to 5 alternate parses, with assumption that both segmentation and line-of-sight is correct.

Alternatives Example:

Written:



Parsed:

$$x^{2-3}$$

Alternatives:

- ① x^{2-3}
- ② $x^2 - 3$
- ③ $x^z - 3$
- ④ x^{2-z}

Alternatives: Symbols

CNN assigns a probability for each node to be each symbol

Example: $z_i^2 = 3$

Symbol Probabilities:

- 1 (z, 17.46), (t, 10.07), (E, 7.00), (e, 6.36), (Σ , 5.94)
- 2 (2, 10.30), (R, 9.89), (a, 6.31), (k, 6.04), (B, 3.85)
- 3 (i, 19.68), (1, 9.96), (j, 7.17), (, , 6.63), (!, 6.20)
- 4 (=, 16.71), (9, 8.21), (5, 7.07), (\div , 6.18), (-, 6.08)
- 5 (3, 10.46), (}, 6.27), (S, 6.08), (B, 5.88), (5, 5.68)

Alternatives: Symbols

How many symbol options to give?

Alternatives: Symbols

How many symbol options to give?

- Provide at least 1 and at most 5 alternatives
- Employ a Support Vector Machine (SVM) to determine the exact number to be displayed

Alternatives: Support Vector Machine

- Tested many possible classifiers: linear SVM resulted in highest recall while maintaining good precision
- Trained on a subset of expressions from CROHME 2019 dataset
- Implemented with the following data points:
(x,y) = (Chosen symbol probability, Alternate symbol probability)

Alternatives: Relations

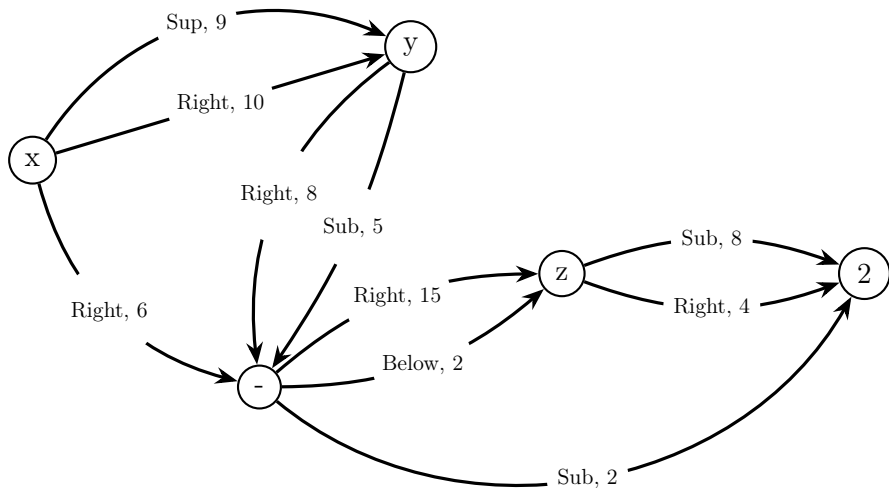
- Recall: Relations between symbols are defined by edges in the graph representation
- Determining alternative relations corresponds to finding the k-best maximum directed rooted spanning trees
- Applies a Beam Search Algorithm as a result of its efficiency and relative accuracy

Alternatives: Beam Search Algorithm

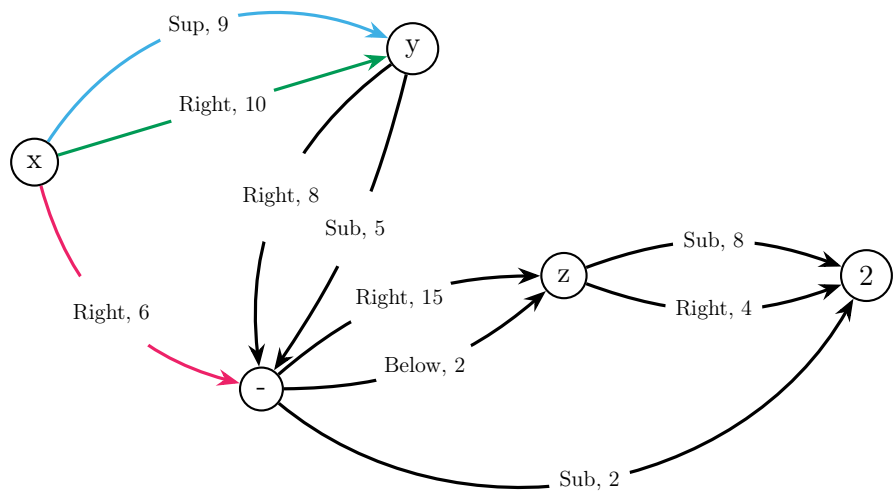
Steps in Beam Search Algorithm:

- ➊ Add root node to list of alternatives
- ➋ Generate all possibilities, for each alternative, that are of the form:
alternative + edge from node in alternative to a node not in alternative
- ➌ Choose the 5 best of all possibilities. These are now the alternatives
- ➍ Repeat steps 2 and 3 until all nodes are included in the alternatives

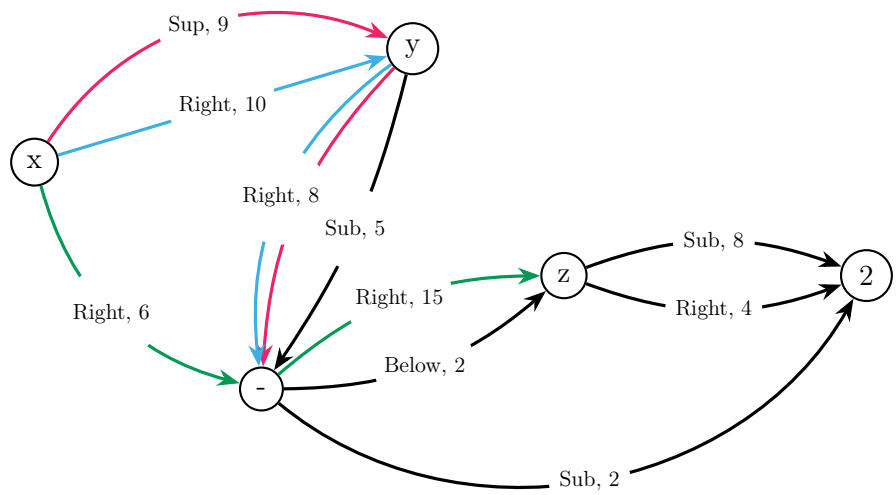
Alternatives: Beam Search Example



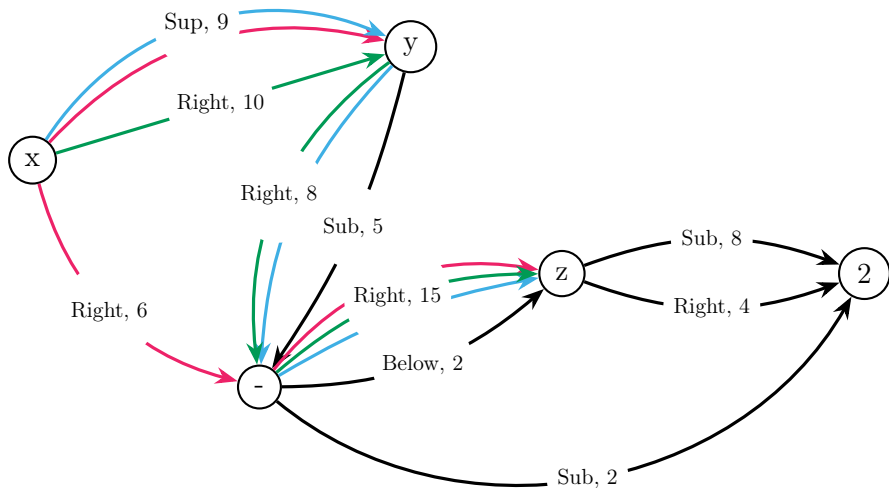
Alternatives: Beam Search Example



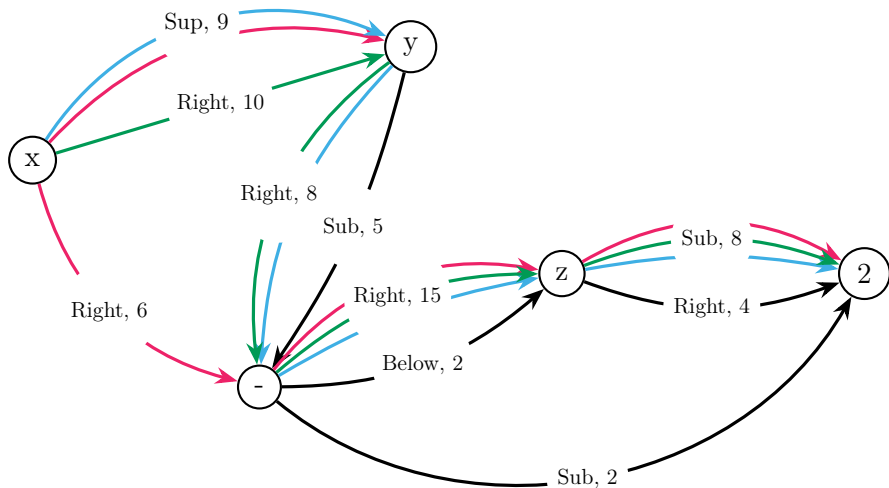
Alternatives: Beam Search Example



Alternatives: Beam Search Example



Alternatives: Beam Search Example



Alternatives: Symbols and Relations

How to have alternatives with a mix of symbol and relation changes?

Alternatives: Symbols and Relations

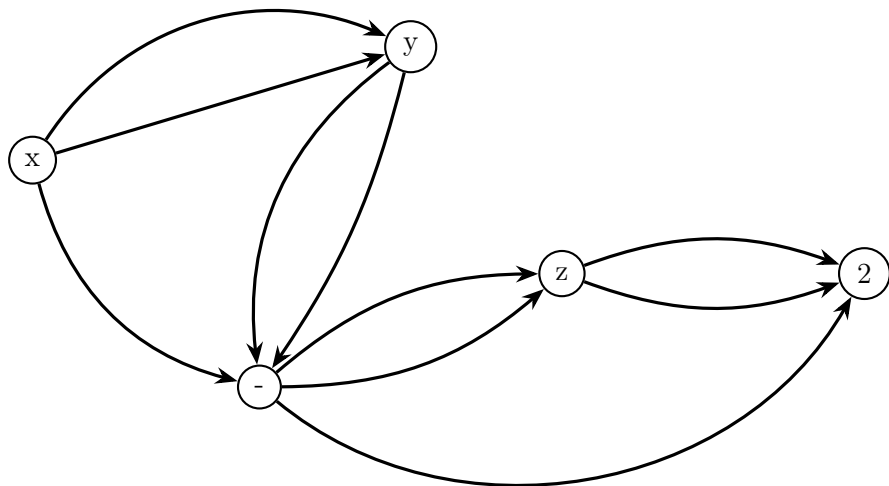
How to have alternatives with a mix of symbol and relation changes?

First, each node is divided into two new nodes:

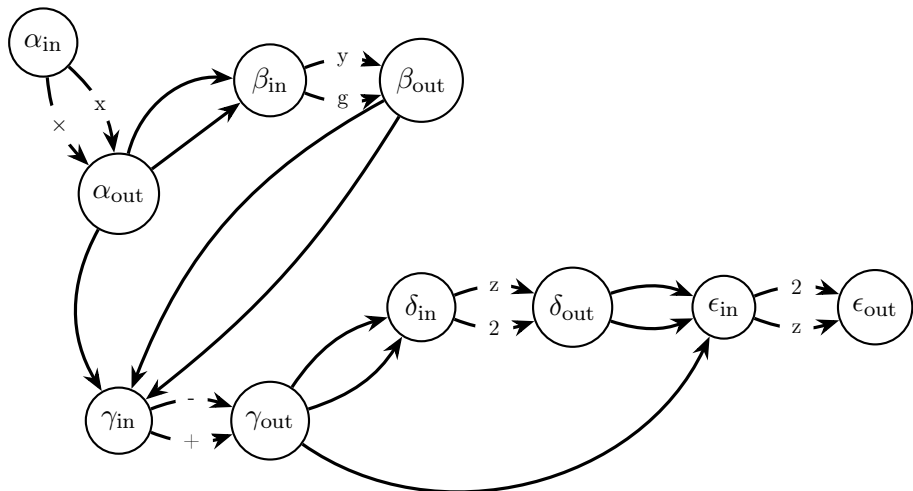
- One has all edges incoming to the node
- One has all edges outgoing from the node
- Edges between the nodes represent symbol alternates
- Probabilities are rescaled to match with relation probabilities probabilities

Then, Beam Search is employed on the new graph, with the edge chosen between divided nodes used as the resultant symbol.

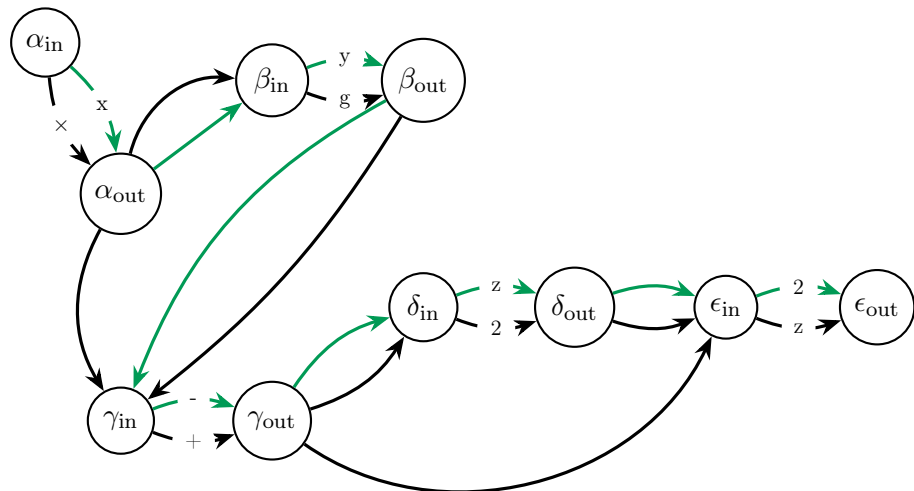
Alternatives: Symbols and Relations Example



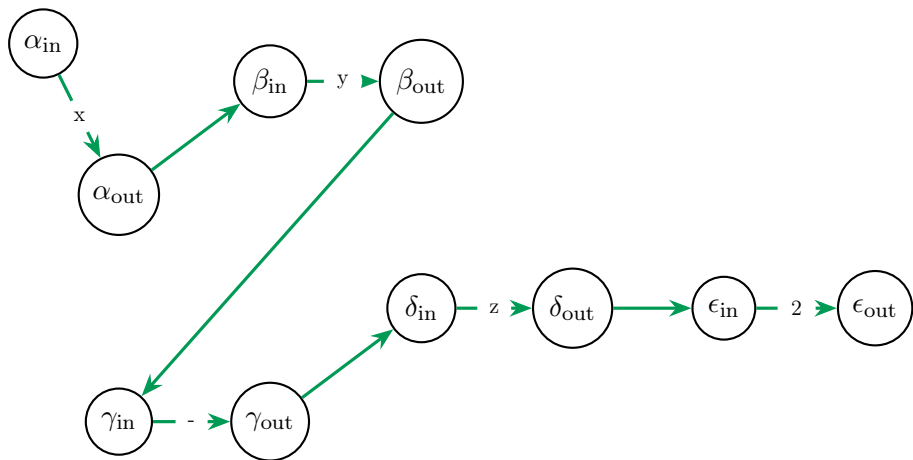
Alternatives: Symbols and Relations Example



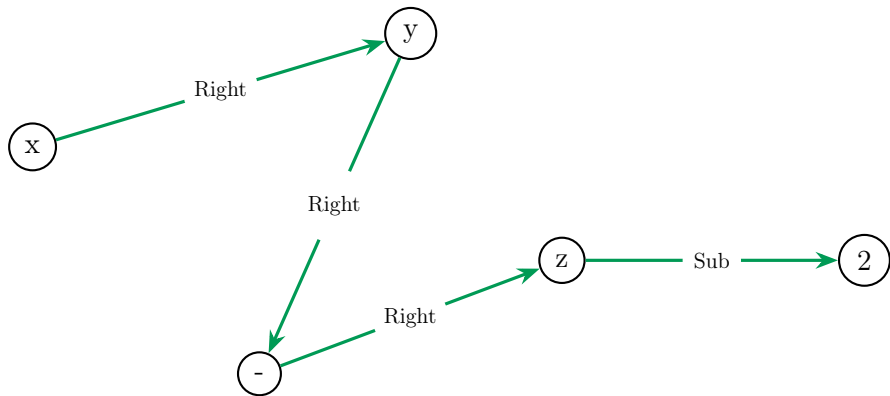
Alternatives: Symbols and Relations Example



Alternatives: Symbols and Relations Example



Alternatives: Symbols and Relations Example



Alternatives: Implementation Summary

Given a graph and a list of nodes recognizer gets alternatives as follows:

Alternatives: Implementation Summary

Given a graph and a list of nodes recognizer gets alternatives as follows:

- If it is a single node, produce alternatives list created with the SVM

Alternatives: Implementation Summary

Given a graph and a list of nodes recognizer gets alternatives as follows:

- If it is a single node, produce alternatives list created with the SVM
- Otherwise:
 - Break the nodes in the graph in two
 - Employ Beam Search
 - Recombine nodes, using edge between as symbol

Alternatives

If we allow alternatives to be used for QD-GGA-ALT, we get the following on the CROHME 2019 Test Set:

Recognizer	Fully Correct
USTC-iFLYTEK	80.73
Samsung R & D 1	79.83
MyScript	79.15
Sun Yat-Sen U.	77.40
QD-GGA-ALT	65.97
Samsung R & D 2	65.97
PAL-v2	62.55
MathType	60.13
QD-GGA	43.30
TUAT	39.95

Alternatives: Test Notes

In comparison to QD-GGA without alternatives:

Alternatives: Test Notes

In comparison to QD-GGA without alternatives:

- Of the originally incorrect recognized expressions with QD-GGA we can rescue **45.74%** to be newly correctly recognized by QD-GGA-ALT

Alternatives: Test Notes

In comparison to QD-GGA without alternatives:

- Of the originally incorrect recognized expressions with QD-GGA we can rescue **45.74%** to be newly correctly recognized by QD-GGA-ALT
- Of these newly correctly recognized expressions, the percentage only requiring one swap for correctness: **71.04%**

Alternatives: Test Notes

In comparison to QD-GGA without alternatives:

- Of the originally incorrect recognized expressions with QD-GGA we can rescue **45.74%** to be newly correctly recognized by QD-GGA-ALT
- Of these newly correctly recognized expressions, the percentage only requiring one swap for correctness: **71.04%**
- Fully Correct Rate: **+22.57%**

QD-GGA Recognizer: Development Roadmap

QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct

QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct
- Starting with the DPRL implementation, incorporate symbols not from the CROHME competition

QD-GGA Recognizer: Symbols

- **Numbers:**

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

- **Lower Case Letters:**

a, b, c, d, e, f, g, h, i, j, k, l, m,
n, o, p, q, r, s, t, u, v, w, x, y, z

- **Upper Case Letters:**

A, B, C, E, F, G, H, I, L,
M, N, P, R, S, T, V, X, Y

- **Greek Letters:**

$\alpha, \beta, \gamma, \Delta, \theta, \lambda, \mu, \pi, \sigma, \phi$

- **Others:**

(,), [,], {, }, +, -, _, !, , , ., /, |, ×, ÷, <, ≤, =, ≥, >, ≠, ±, sin,
cos, tan, ∫, ∑, √, lim, log, ∞, †, ∃, ∀, ∈, ..., →

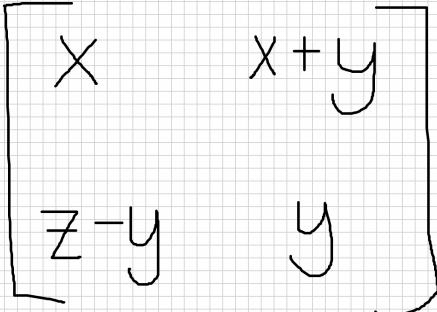
QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct
- Starting with the DPRL implementation, incorporate symbols not from the CROHME competition
- Support matrices and multiple expressions

Matrices

BrushSearch Math_Stack_Exchange File Settings

```
\begin{bmatrix} x & x + y \\ z - y & y \end{bmatrix}
```

$$\begin{bmatrix} x & x + y \\ z - y & y \end{bmatrix}$$


A hand-drawn matrix on a grid background, enclosed in large square brackets. The matrix is a 2x2 grid with the following elements:

- Top-left: x
- Top-right: $x + y$
- Bottom-left: $z - y$
- Bottom-right: y

Multiple Expressions

The image shows a screenshot of the BrushSearch application. The top bar is blue and contains the text "BrushSearch" followed by three dropdown menus: "Math_Stack_Exchange", "File", and "Settings". Below the top bar is a code editor with a white background and a vertical scrollbar on the right. The code editor contains the text $x + y = 7$ followed by a double backslash $\backslash\backslash$ and then $3x - y = 1$. To the right of the code editor, there is a preview area with a white background and a vertical scrollbar on the left. The preview area contains the text $x + y = 7$ followed by $3x - y = 1$. Below the code editor and preview area is a drawing canvas with a light blue grid background. The canvas contains two handwritten equations: $x + y = 7$ and $3x - y = 1$. On the left side of the drawing canvas, there is a vertical toolbar with several icons: a trash can, a circular arrow, a pencil, an eraser, a hand, a checkmark, a circular arrow, a download icon, and a blue circle.

QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct
- Starting with the DPRL implementation, incorporate symbols not from the CROHME competition
- Support matrices and multiple expressions
- Allow the user to train the recognizer

Training

The screenshot displays the BrushSearch application interface. In the background, a math editor shows the LaTeX expression $\int_0^{\infty} \frac{1}{\sqrt{x^3 - 2}}$. The foreground features a 'Profile' configuration window for the user 'mirettemarzouk@gmail.com'. The window includes a 'Save Profile' button, a 'Discard Changes' button, and a checkbox for 'Add my profile during recognition'. Below these are tabs for 'Numbers', 'Lower Case', 'Upper Case', 'Greek Symbols', and 'Others'. The 'Numbers' tab is active, showing a list of numbers 0 through 7. The number '1' is selected, and three handwritten examples of the digit '1' are displayed in a grid. Each example has a trash icon below it, and a plus sign icon is located below the first example.

QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct
- Starting with the DPRL implementation, incorporate symbols not from the CROHME competition
- Support matrices and multiple expressions
- Allow the user to train the recognizer
- Use user's feedback to improve recognition

QD-GGA Recognizer: Development Roadmap

- Within alternatives, don't assume that the segmentation and line-of-sight is correct
- Starting with the DPRL implementation, incorporate symbols not from the CROHME competition
- Support matrices and multiple expressions
- Allow the user to train the recognizer
- Use user's feedback to improve recognition

Thank you!

References

- 1 Mahdavi, Mahshad, "Query-Driven Global Graph Attention Model for Visual Parsing: Recognizing Handwritten and Typeset Math Formulas" (2020). Thesis. Rochester Institute of Technology.
- 2 Mahdavi, M., Sun, L., Zanibbi, R. (2020). Visual Parsing with Query- Driven Global Graph Attention (QD-GGA): Preliminary Results for Handwritten Math Formula Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 570-571).
- 3 Mahdavi, M., Zanibbi, R., Mouchère, H., Viard-Gaudin, C., Garain, U. (2019, September). ICDAR 2019 CROHME+ TFD: Competition on recognition of handwritten mathematical expressions and typeset formula detection. In 2019 International Conference on Document Analysis and Recognition (ICDAR) (pp. 1533-1538). IEEE.
- 4 Zanibbi, R., Mouchère, H., and Viard-Gaudin, C. "Evaluating structural pattern recognition for handwritten math via primitive label graphs," Proc. Document Recognition and Retrieval XX, San Francisco, CA, Feb. 2012.