

CROHME 2014

Competition on Recognition of
Online Handwritten Mathematical Expressions

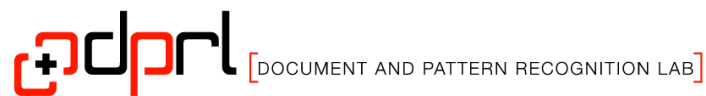
H. Mouchère
C. Viard-Gaudin



UNIVERSITÉ DE NANTES



R. Zanibbi

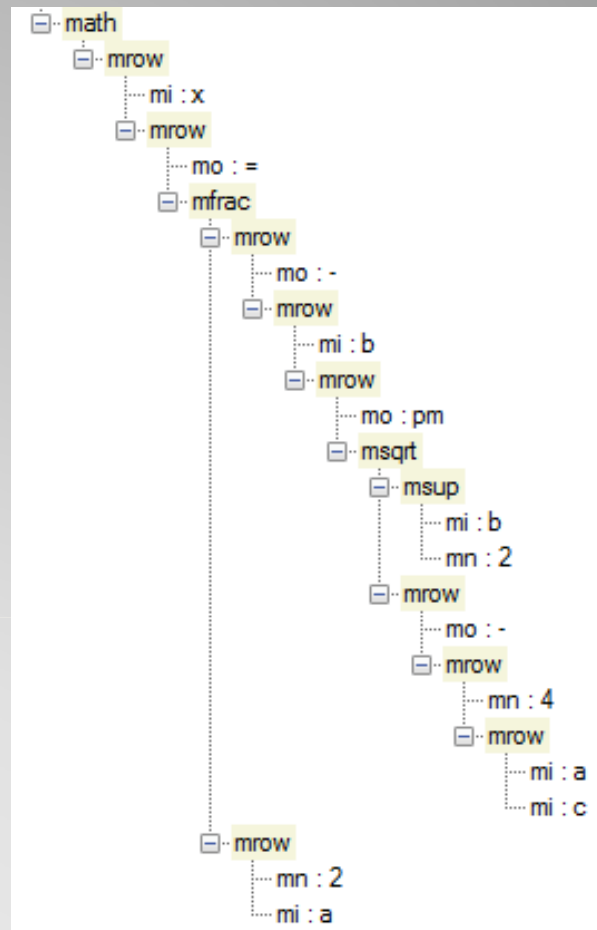
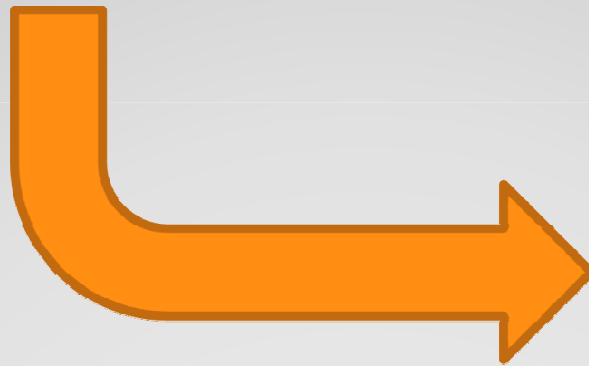


U. Garain



- Online Handwritten Math Expression

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$



From input to output

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

CROHME 2014 4th edition

3 Tasks:

- Isolated symbols (Task 1)
- Full expressions (Task 2)
- Matrices (Task 3)

Data collection

File format

Semi manual
labeling

Run systems

Evaluation
tools

Sharing data

New test sets

Math ML

≠ config.

Scripts

Package
(USB, TC 11)

- CROHME Evolution

	Beijing	Bari	Washington	Creta
	2011	2012	2013	2014
Expression Grammars	2	3	2	1
# Symbol Classes	60	75	101	101
<i>Task 2</i> Training expressions	921	1 341	8 836	8 836
<i>Task 2</i> Test expressions	348	486	671	986
Participations for <i>Task 1</i>	-	-	-	8
Participations for <i>Task 2</i>	4	6	8	7
Participations for <i>Task 3</i>	-	-	-	2

- CROHME Task 1 (Symbols segmented from Task 2)

Training Set	Valid Symbols	85 781
	Invalid Symb. (Junk)	User-specified (script)
Test Set	Valid Symbols	10 061
	Invalid Symb. (Junk)	9 161

- CROHME Task 3

Training Set	Matrices	362
	Number of cells	2 332
	Number of symbols	4 281
Test Set	Matrices	175
	Number of cells	1 075
	Number of symbols	2 101

- CROHME Participants

System	Affiliation	Tasks
I	Universitat Politècnica de València	1 – 2 – 3
II	University of São Paulo	1 – 2
III	MyScript	1 – 2 – 3
IV	Rochester Institute of Technology, DRPL	1 – 2
V	Rochester Institute of Technology, CIS	1 – 2
VI	Tokyo University of Agriculture and Technology	1 – 2
VII	University of Nantes, IRCCyN	2
VIII	ILSP/Athena Research and Innovation Center	1

- CROHME Results

System	Without Junk (101 classes)		With Junk			
	Top1	True Mean Position	(102 classes)		(2 classes)	
			Top1	True Mean Position	TAR	FAR
I	91.24	1.20	n.r.o.			
	89.79	1.25	84.14	1.28	80.29	6.44
II	82.72	1.60	79.11	1.50	82.49	14.63
III	91.04	1.19	85.54	1.23	87.12	10.39
IV	88.66	1.27	83.61	1.29	83.52	9.03
V	85.00	1.37	71.19	1.48	86.84	36.85
VI	84.31	1.53	n.r.o.			
	82.08	1.63	76.24	1.57	77.98	16.47
VIII	77.25	1.68	n.r.o.			

**Isolated Symbol recognition
(TASK 1)**

- CROHME Results

	System I	System III
Expression Rate	31.15	53.28
Symbol Reco Rate	87.43	89.81
Matrix Recall Rate	73.14	92.57
Row Recall Rate	70.59	92.00
Column Recall Rate	50.84	69.16
Cell Recall Rate	55.35	71.07

Matrix Recognition (TASK 3)

- CROHME Results

System	Correc (%)t	<= 1 error	<= 2 errors	<= 3 errors
I	37.22	44.22	47.26	50.20
II	15.01	22.31	26.57	27.69
III	62.68	72.31	75.15	76.88
IV	18.97	28.19	32.35	33.37
V	18.97	26.37	30.83	32.96
VI	25.66	33.16	35.90	37.32
VII	26.06	33.87	38.54	39.96

**Full Expression Recognition
(TASK 2)**

- CROHME Winners
For Tasks 1, 2 and 3:
 - Best system:
MyScript, France (System-III)
 - Best system trained on CROHME data:
Univ. of Valencia, Spain (System-I)