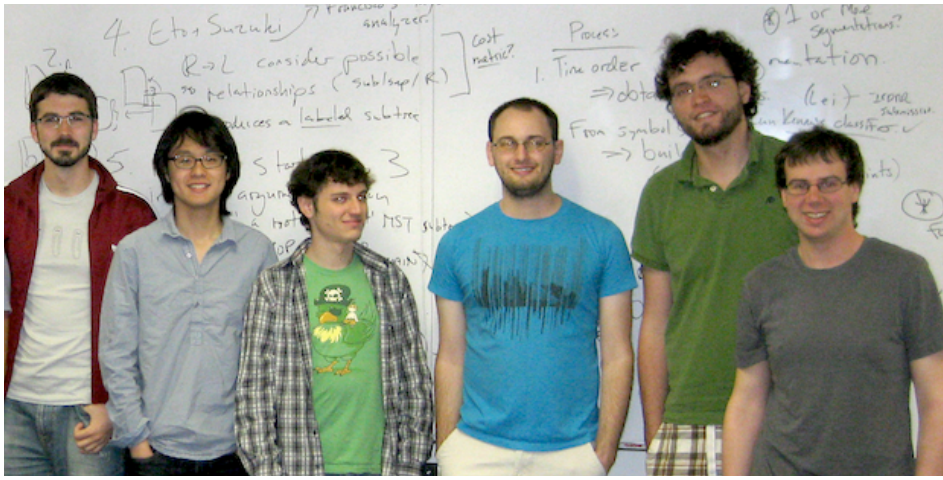
dprl [DOCUMENT AND PATTERN RECOGNITION LAB]



dprl@10:
The Document and Pattern Recognition Lab's First 10 Years

Dr. Richard Zanibbi
Director, dprl

Goal

Improve algorithms and tools for recognition and retrieval of information in documents, images, audio and video.

Let's first look at some examples of systems and results produced by student researchers in the dprl.

Video CAPTCHA
(Kurt Kluever, CUPS 2009)

- Identify computers ('bots') vs. humans. One 'correct' word to pass. Words stemmed (via Porter), synonyms allowed.
- 'Correct' answers generated automatically from tags on 'similar' YouTube videos (by cosine similarity). Filter tags with frequency > t %.
- In an experiment, students passed 90% of challenges, vs. 13% break rate using most common tags.
- **Innovations:** First video CAPTCHA. Random walk-based video sampling, tag set expansion and filtering methods.

Text-Conv: Text Detection in Natural Scenes

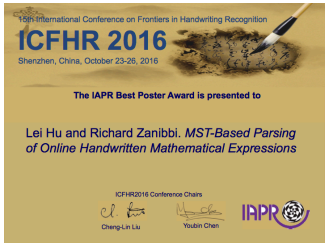
(Siyu Zhu, CVPR 2016; David Syner, Bo Ding, Kardo Aziz)

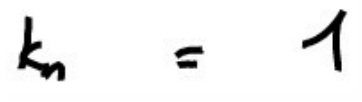
- Cascaded AdaBoost-based detector that obtained state-of-the-art results for ICDAR 2015 benchmark without OCR or conv. nets.
- **Innovations**: patch context (8-neighborhood); sampling techniques (coarse to fine; graph constraints); MST word segmentation; extended convolutional k-means [Coates & Ng] for patch learning.

(a) Input (b) Local 3x3 (c) Contextual 9x9

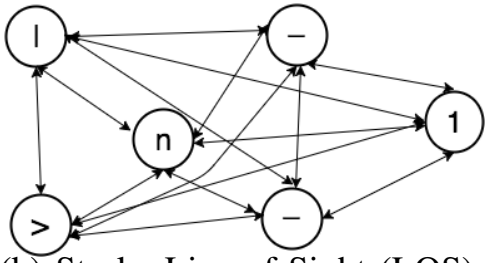
Handwritten Math Recognition

(Lei Hu, ICFHR 2016; Ouyang, Davila, Condon, Ravi)

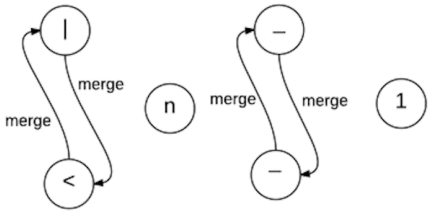




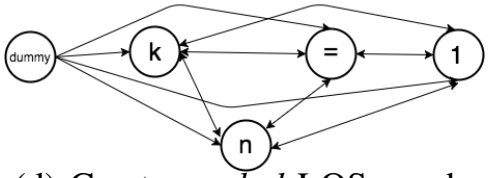
(a) **Input:** Six strokes for $k_n = 1$



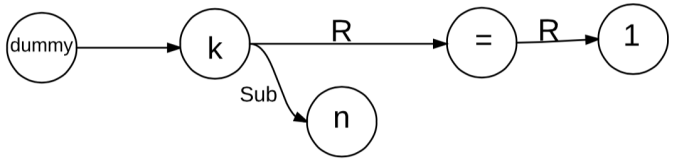
(b) Stroke Line-of-Sight (LOS) graph



(c) Segment symbols w. binary directed edge classifier [15]

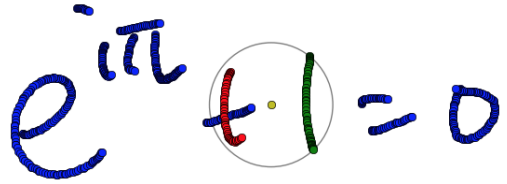


(d) Create *symbol* LOS graph. Classify symbols & compute relationship probabilities

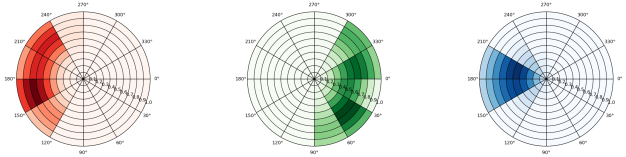


(e) Apply Edmonds' algorithm to obtain MST.
Output: Remove dummy node for Symbol Layout Tree

- Efficient parser for math written on a tablet.
- **Innovations:** LOS constraints; Parzen shape context features; accurate segmentation & structure without OCR



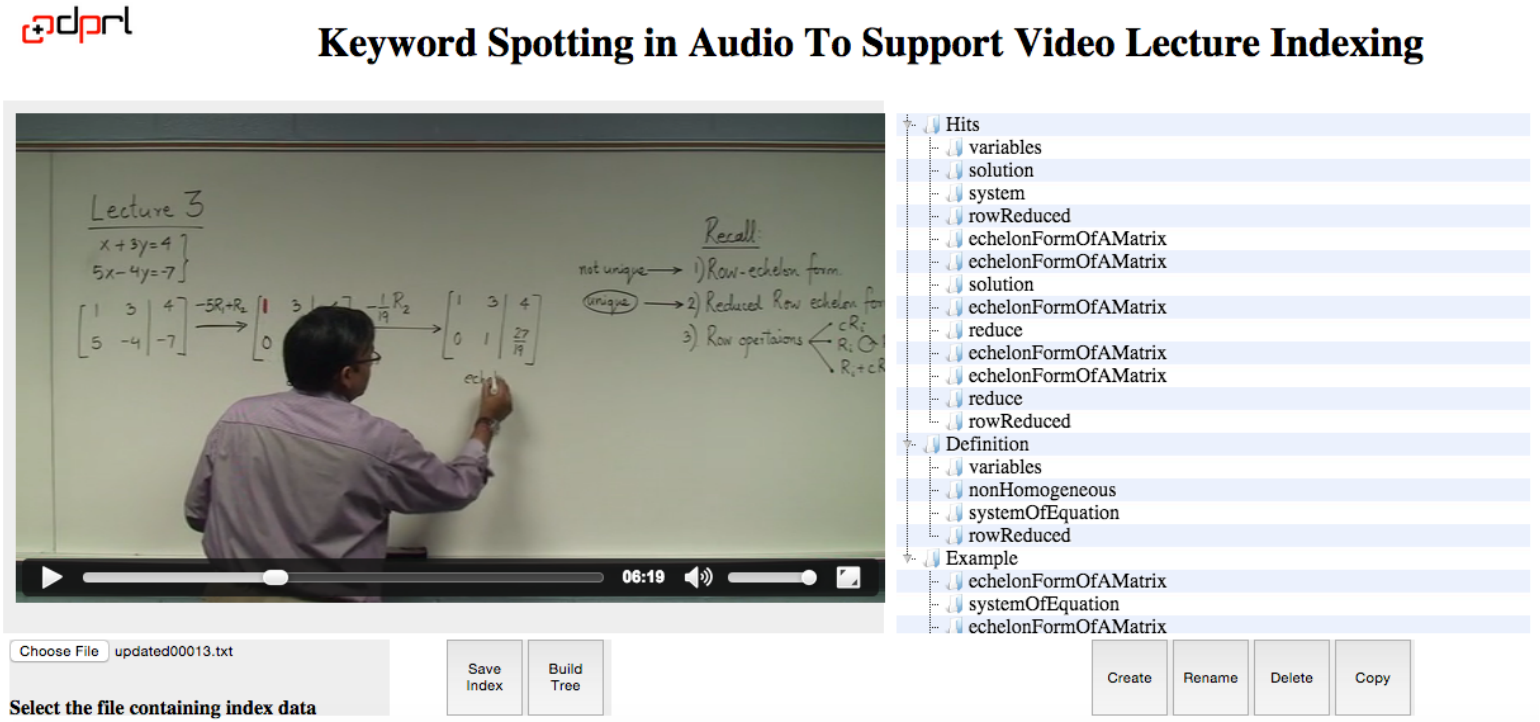
Expression with PSC center and perimeter shown



Parent stroke | Child stroke | Other strokes

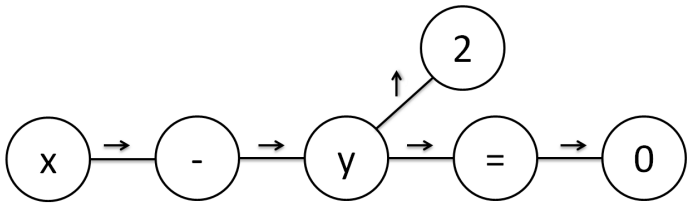
Keyword Spotting in Lecture Audio

(Miller, Kanadje, Agarwal, et al., PR Letters, 2016; Miller-Jacobson)

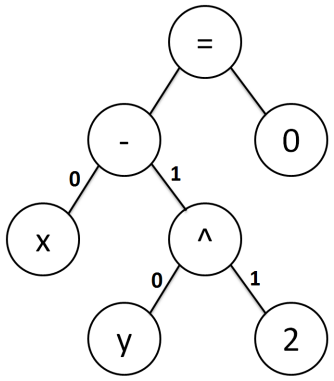


- Instructor searches and indexes lecture videos with queries recorded on a laptop. Interface to categorize and organize contents.
- Unsupervised MFCC + Segmental DTW [Park & Glass]. Average 70% of top 10 hits were correct; within lecture queries 80%; 90% for lapel mic-recorded MIT lectures.
- **Innovations:** whitening of MFCC vectors, reducing strength of low frequencies on laptop recordings, noise in audio recording.

Tangent Formula Search Engine
(Davila, SIGIR 2016 & 2017; Schellenberg, Stalnaker, Pattaniyil)



(a) Symbol Layout Tree



(b) Operator Tree

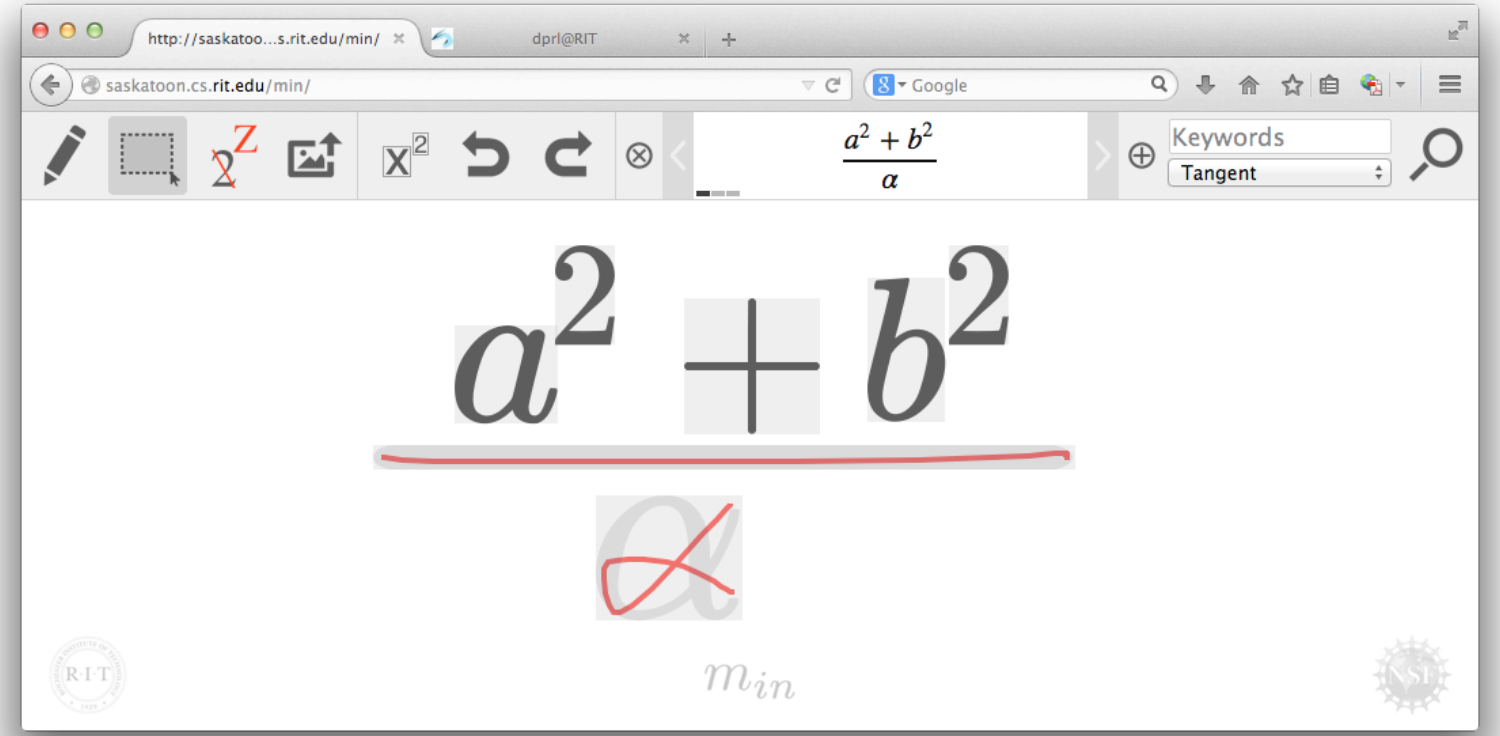
Figure 1: Tree representations for $x - y^2 = 0$

	Core	Matching
1.	$\mathcal{O}(mn \log m)$	$\mathcal{O}(mn \log m)$
2.	$\mathcal{O}(mn)$	$\mathcal{O}(VE \log V)$
3.	$\mathcal{O}(mnp)$	$\mathcal{O}(VE \log V \log(VC))$
4.	$\mathcal{O}(m + \log n)$	$\mathcal{O}(Tm) = \mathcal{O}\left(n^2 m \log n\right)$
5.	$\mathcal{O}\left(m\sqrt{n} \log n\right)$	$\mathcal{O}(nk \log k)$

- Symbol pair path matching, re-ranking by best query match.
- **Innovations:** symbol pair-based model. Unification and wildcard support in rerank. Near state-of-art results for NTCIR 12 in Tokyo; orders of magnitude faster than ‘best’ system (real-time)

m_{in} Math Search Interface

(Sasarak, Hart, Pospesel, Stalnaker, Hu, LiVolsi, Zhu, HCIR 2012; Wangari, SIGIR 2014; Orakwue, CICM 2015)



- **Innovations**: first math-aware web search interface with multi-modal math input (drawing, keyboard (LaTeX support), images)

Where do the dprl students come from?

Interdisciplinary by accident (i.e., due to need or synergies), not by design.

- Computer Science
- Interactive Games and Media
- Human-Computer Interaction
- Imaging Science
- Computer Engineering
- Visitors: CS PhD students from Spain, Japan, Tunisia

- All levels: Undergraduate, Master's and PhD students

Where do dprl students go after graduation?

- Google (x7) - including UX research lab at Mountain View
- PhD, IIT (Illinois Institute of Technology); PhD at Rice University; PhD admission offers at Univ. Waterloo (Canada), Univ. Melbourne (Australia)
- Buckler Lab, Cornell University (research programmer in genetics lab)
- Apple (research team that contributed to handwriting rec. for Apple Watch)
- Amazon
- Square
- AppNexus
- …

The dprl as *Student-Centered* Research Lab

- **Assumption:** At its best, research produces helpful knowledge and tools for others. Good papers and good prototypes should be good *tools.*

- Research is used as a pedagogical tool (problem-based learning). Interesting models and results are key, but not the only goals. Others important skills to learn include:
  - oral and written communication,
  - literature review,
  - teamwork,
  - modeling/creative problem solving (want solutions that are effective and elegant)
  - bench science,
  - analysis,
  - finding your own working style and strategies (e.g., coping with getting stuck)

- Often, students work on projects related by subject (e.g., 'math'), or technique(s), but not the same problem. This generates interest and produces a helpful diversity of perspectives, while avoiding a lack of focus and direct competition between students.

The dprl as *Student-Centered* Research Lab

- Replicate, replicate, replicate. Many of the lab's best results came after a student had replicated an earlier result, often having to infer omissions from research papers. This allows students to *feel* a problem and technique directly, with a point of reference, and tends to lead naturally to *targeted* reading.

- Size matters: for myself, 5-7 is ideal, in terms of quality of student instruction, and quality of research output. This seems to lead to 'enough' diversity and discussion for students in the lab to gel as a group, and find a rhythm.

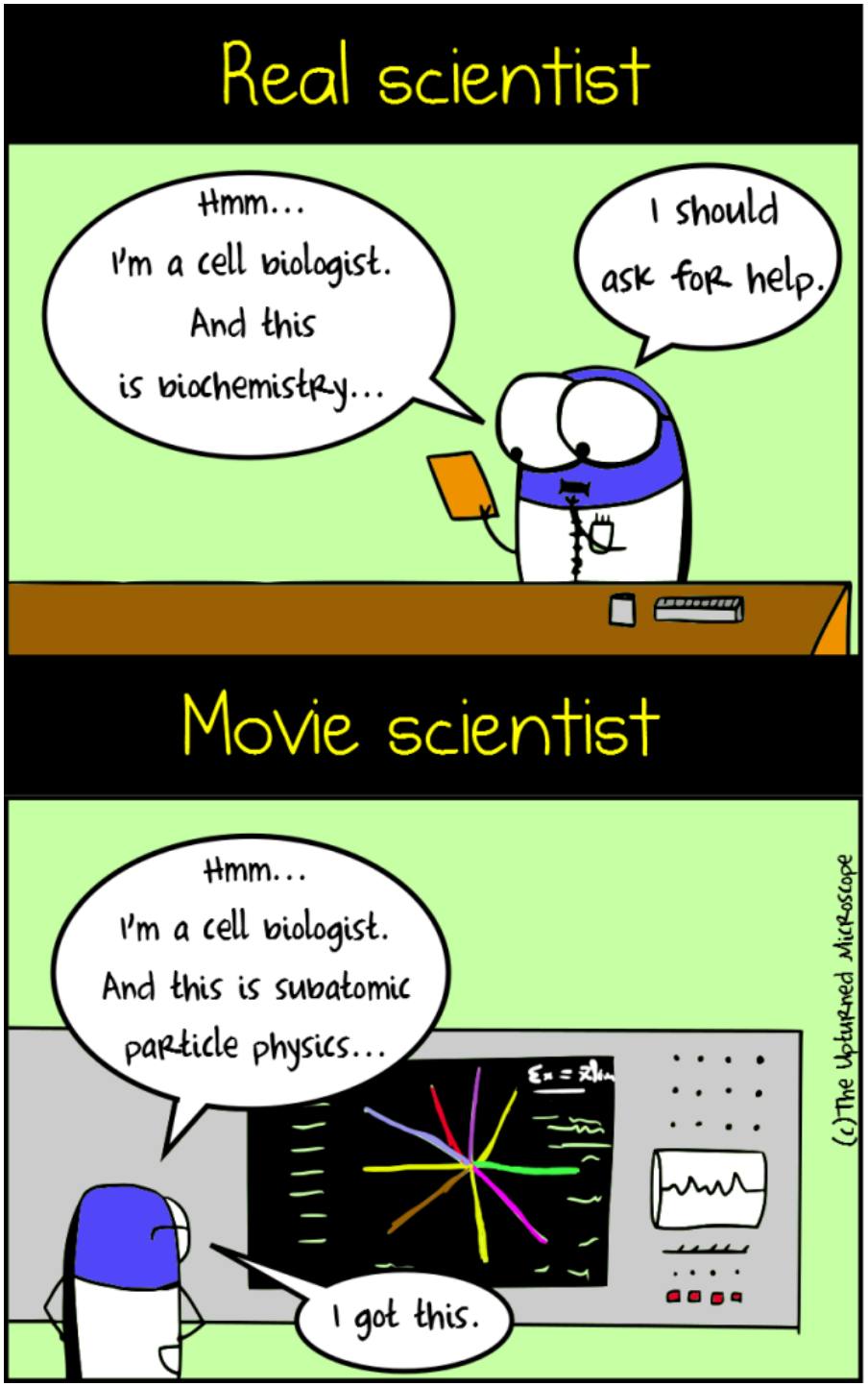
- Meetings matter. The lab holds bi-weekly meetings where students briefly present and discuss their progress with the lab (~10-15 minutes each), in addition to weekly private meetings with myself. Regular discussion is important for progress; equally important is leaving students alone to work comfortably on their own.

- Support learning and progress, not success. Some of the lab's successes were initially thought to be major failures by students. Some students have benefitted from being in the lab, while producing only preliminary or negative results. Happy, healthy people need to find learning and progress, and not just success rewarding.

Thank you.

- We gratefully acknowledge support from the organizations below.

- My sincerest thanks to the dprl students over the last 10 years; it has been a privilege to work with you.


Real scientist
Hmm...
I'm a cell biologist.
And this
is biochemistry...
I should
ask for help.
Movie scientist
Hmm...
I'm a cell biologist.
And this is subatomic
particle physics...
I got this.
(c)The Upturned Microscope