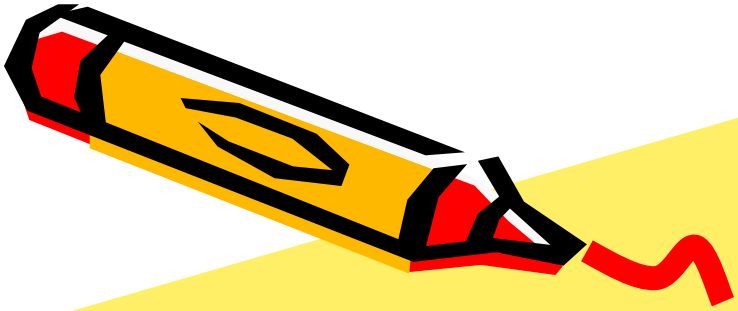
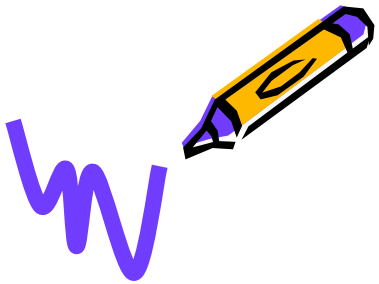


An Efficient Graph-Based Symbol Recognizer

WeeSan Lee <weesan@cs.ucr.edu>
Levent Burak Kara <lkara@andrew.cmu.edu>
Thomas F. Stahovich <stahov@engr.ucr.edu>

http://www.cs.ucr.edu/~weesan/research/graph_matching_slides.pdf

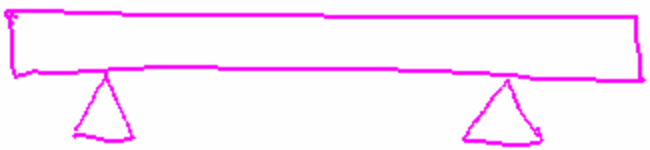


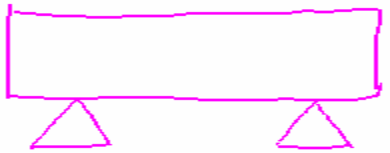
UNIVERSITY of CALIFORNIA
Riverside

Smart Tools Lab

Goals

- Hand-drawn symbol recognizer
- Insensitive to:
  - Uniform / non-uniform scaling
  - Orientation
  - Drawing order
- Efficient
- Easily trainable





Approach

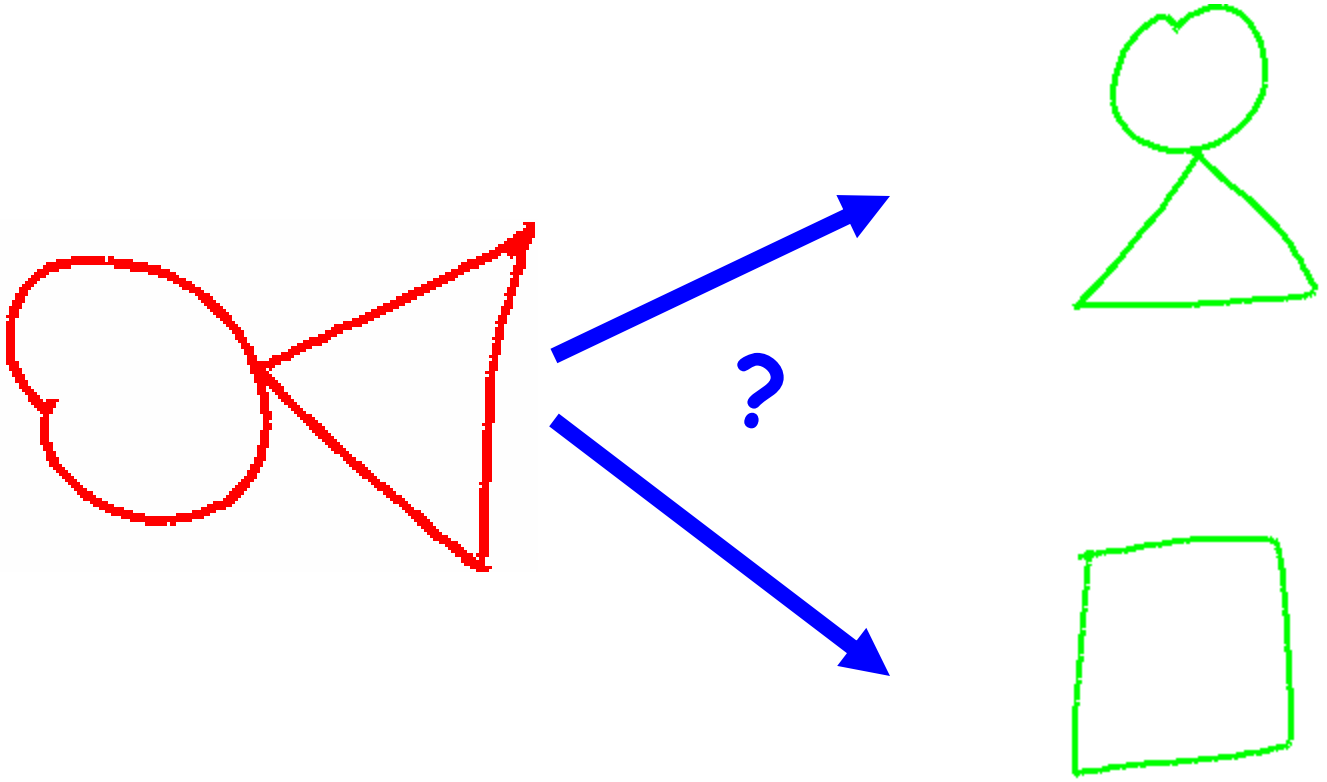
- Representation:
  - Attributed Relational Graph
  - Geometry and topology of a symbol

- Recognition:
  - Approximate graph matching

Recognition

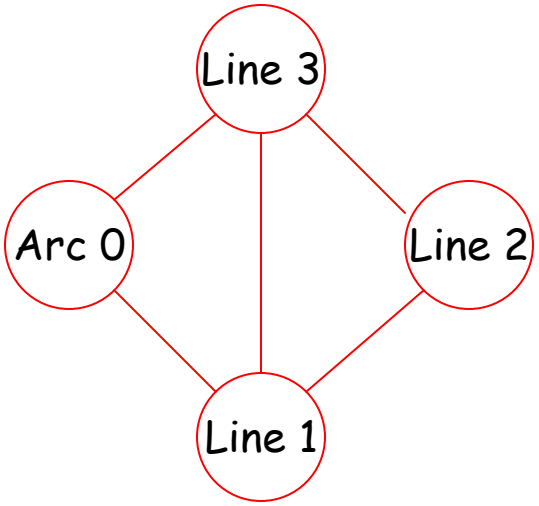
Unknown symbol

Definition symbols



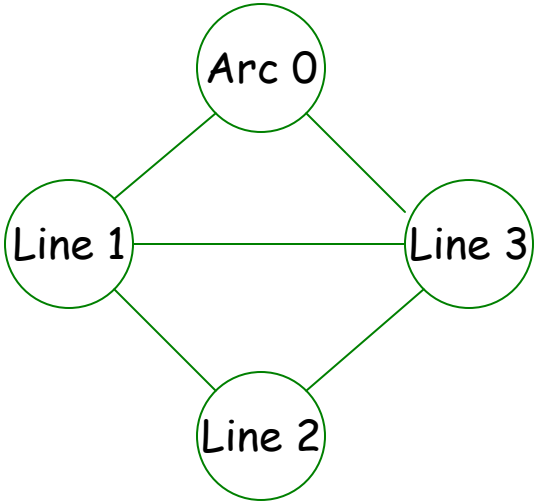
Recognition

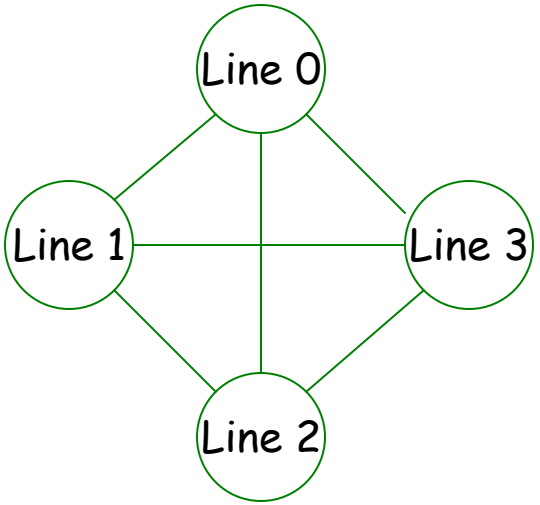
Unknown symbol



?

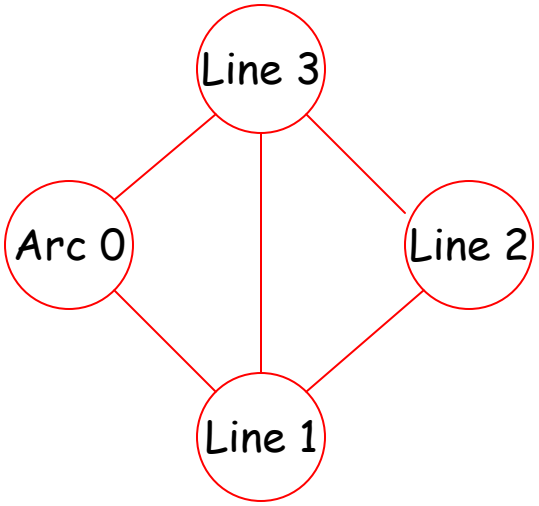
Definition symbols



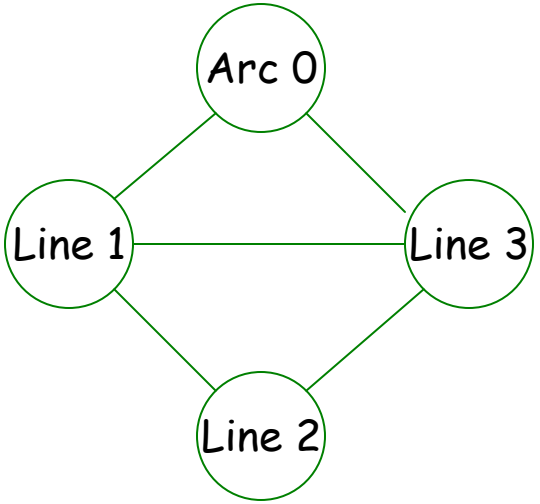


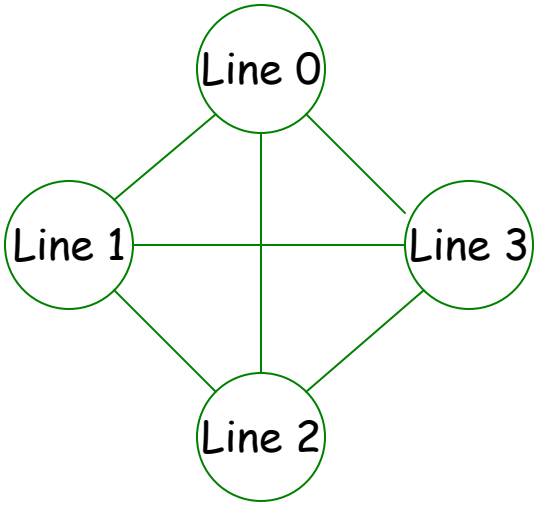
Recognition

Unknown symbol



Definition symbols



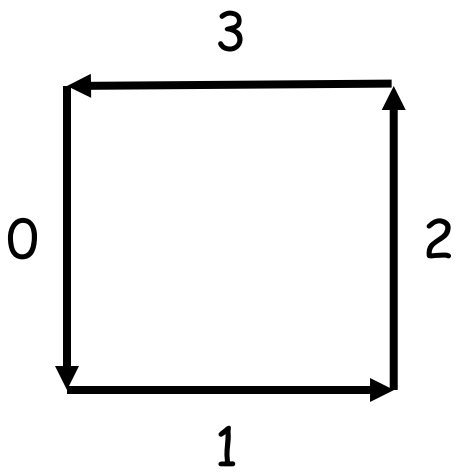


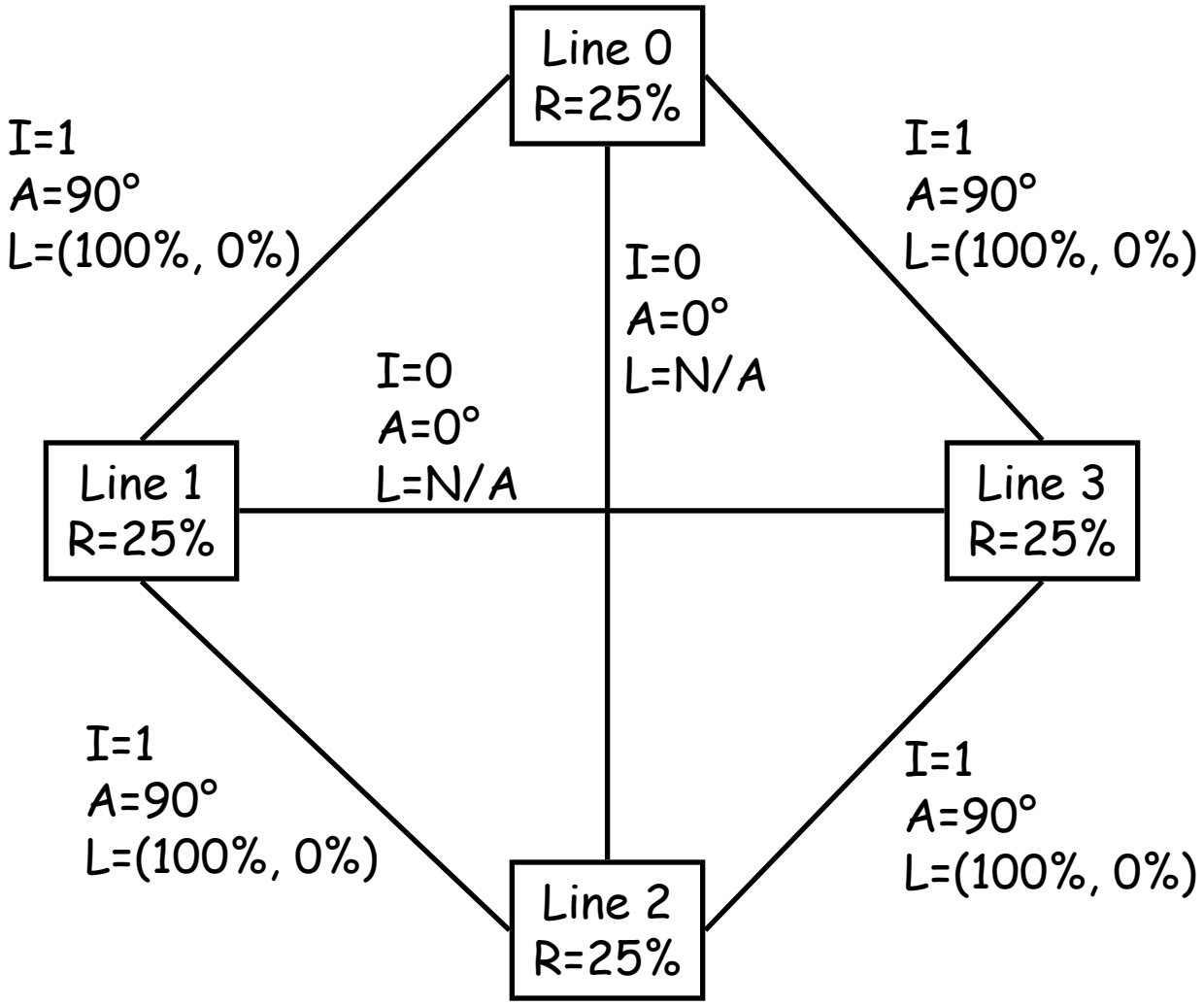
Roadmap

- Representation
- Measuring Similarity
- Graph matching
- User Study
- Conclusions

Representation

Ideal Square

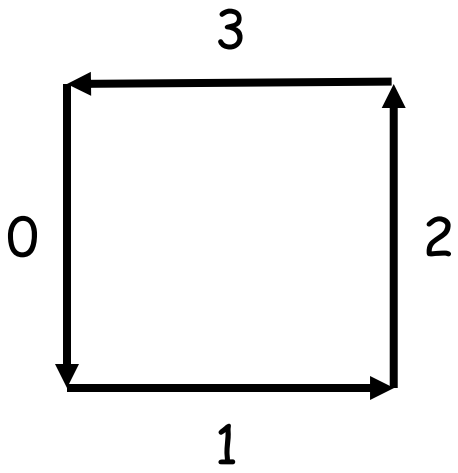




Representation

Line 0
R=25%

Ideal Square



Line 1
R=25%

Line 3
R=25%

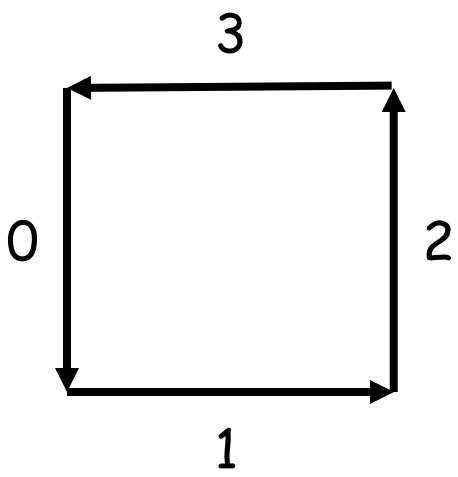
Relative length → scale independence

Large σ → insensitive to *non-uniform* scaling

Line 2
R=25%

Representation

Ideal Square



I=1
A=90°
L=(100%, 0%)

Line 0
R=25%

Line 1
R=25%

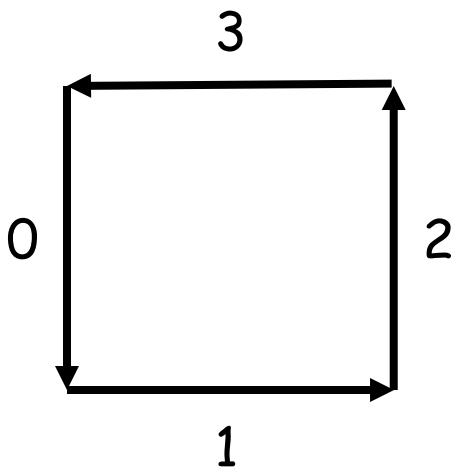
Line 2
R=25%

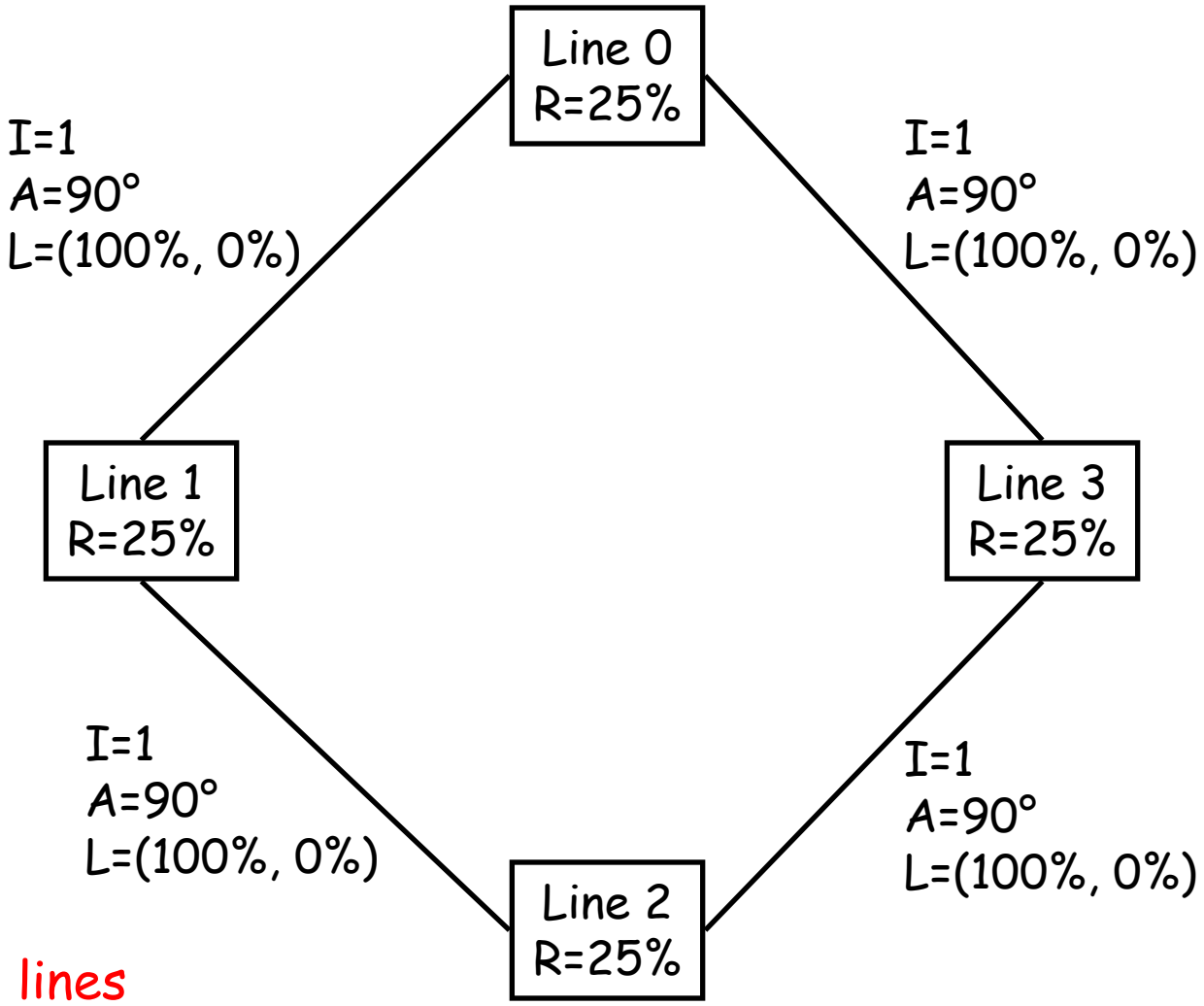
Line 3
R=25%

Intersection Angle:
Defined for *all* pairs of lines

Representation

Ideal Square

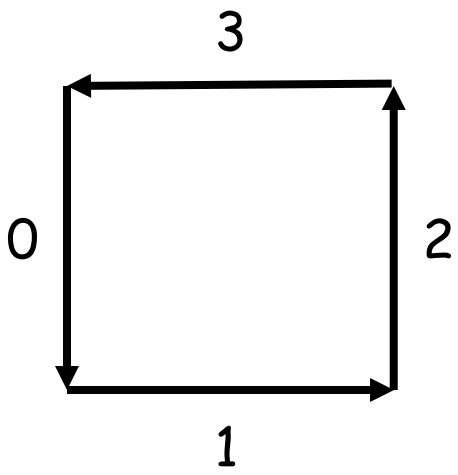


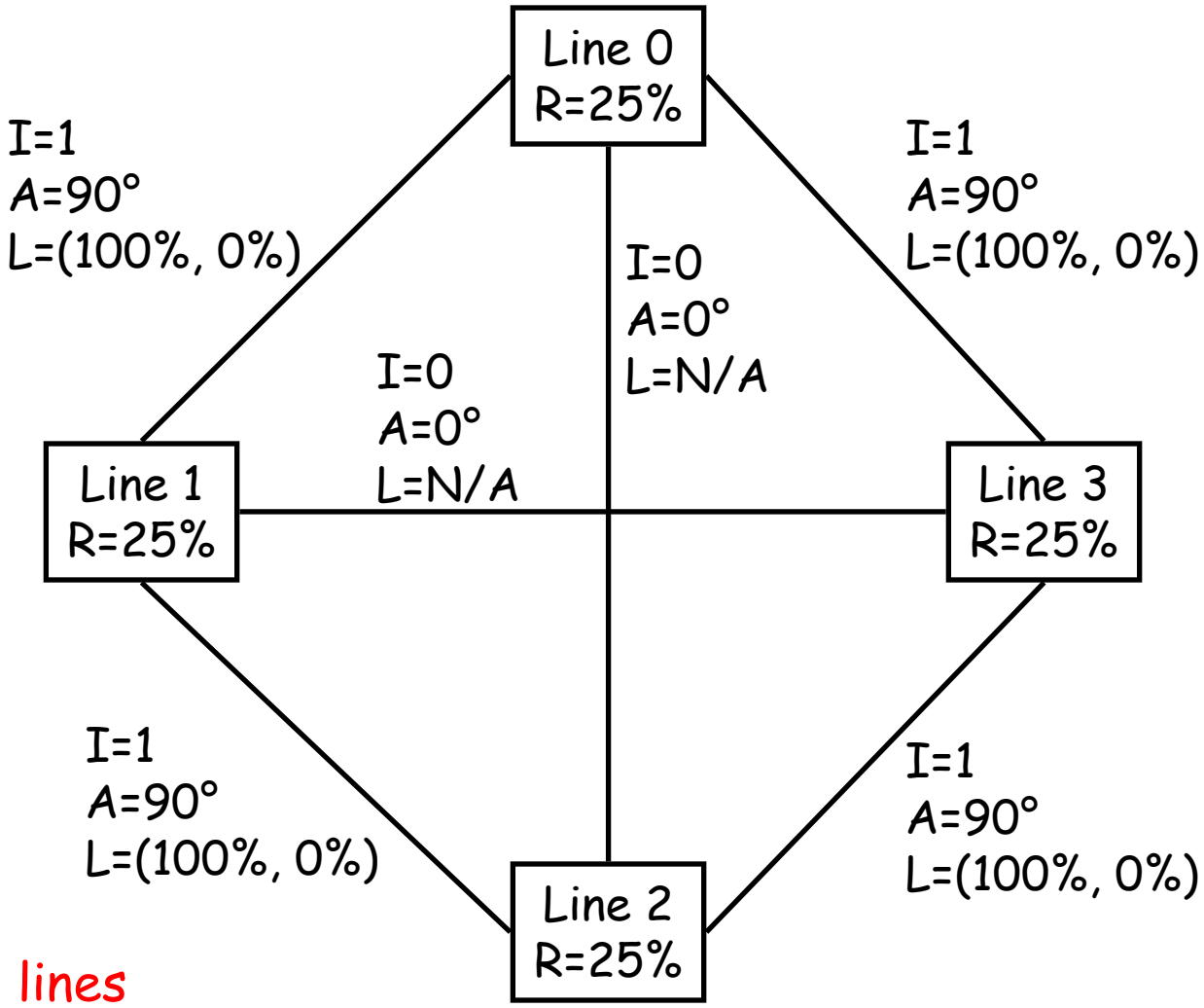


Intersection Angle:
Defined for ***all*** pairs of lines

Representation

Ideal Square





Intersection Angle:
Defined for ***all*** pairs of lines

Representation

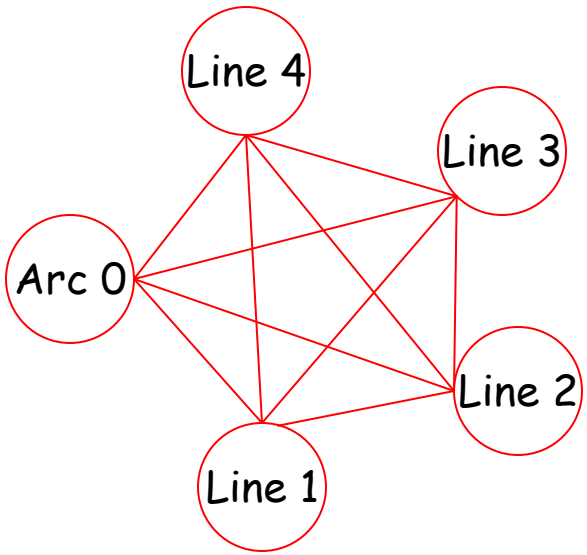
- Definition
  - "Average graph"
  - Statistical model -> μ, σ
- Segmentation
  - Pen strokes -> primitives
    - Lines and arcs
  - Speed-based segmenter

Roadmap

- ✓ Representation
- Measuring Similarity
- Graph matching
- User Study
- Conclusions

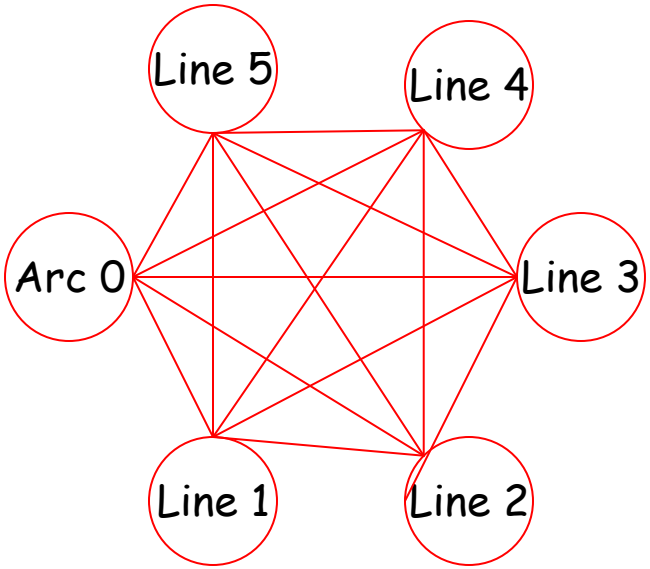
Measuring Similarity

Unknown symbol



?

Definition symbol



Measuring Similarity

$$Similarity\ Score = 1 - \sum_{i=1}^{6} w_i E_i$$

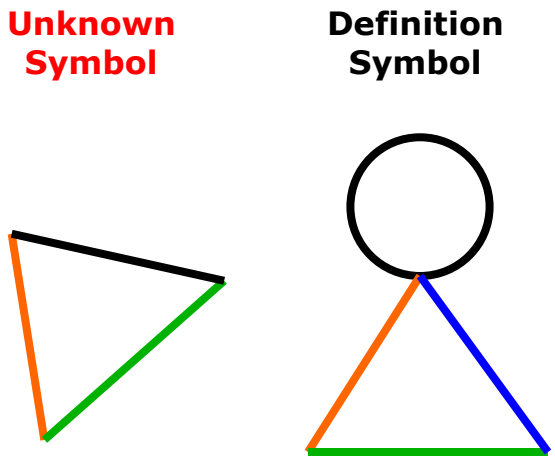
Error Metrics (E_i)	**Weight** (w_i)
E_1: Primitive count error	20%
E_2: Primitive type error	20%
E_3: Relative length error	20%
E_4: Number of intersections error	15%
E_5: Intersection angle error	15%
E_6: Intersection location error	10%

Primitive Count Error

$$E_1 = \frac{difference\_in\_primitive\_count}{MIN\_primitive\_count}$$

Example:

E_1 = 1/3

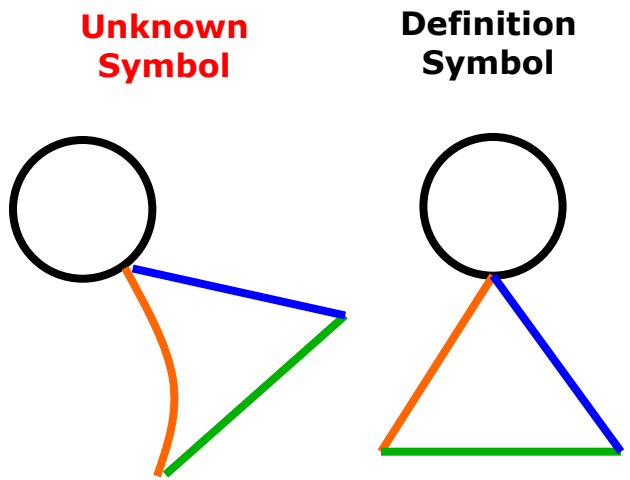


Primitive Type Error

$$E_2 = \frac{\#\_of\_type\_mismatches}{MIN\_primitive\_count}$$

Example:

E_2 = 1/4



Relative Length Error

$$E_3 = \frac{\sum \mathrm{Re}\,lative\_length\_errors}{MIN\_primitive\_count}$$

$$\mathrm{Re}\,lative\_length\_error = 1 - P(R, \mu, \sigma)$$

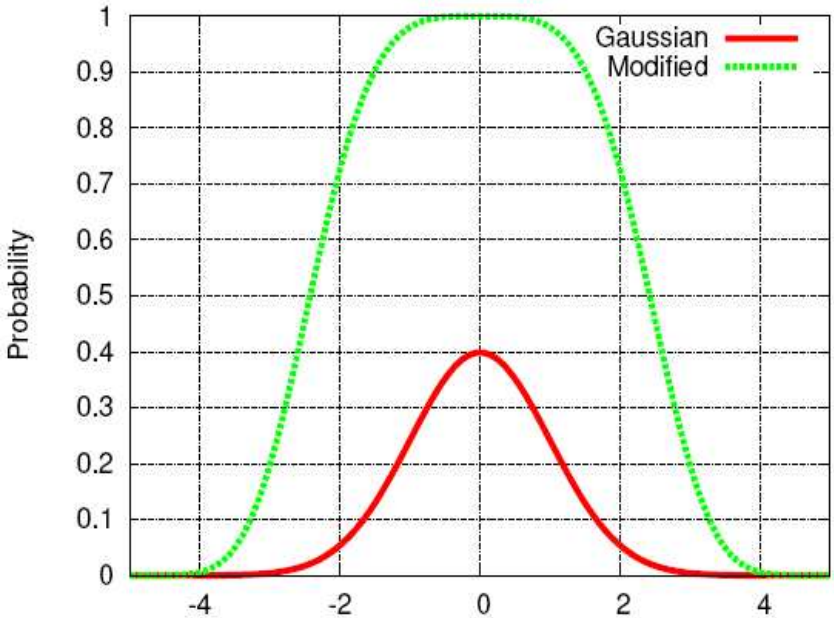
Relative Length Error

$$E_3 = \frac{\sum \mathrm{Re}\,lative\_length\_errors}{MIN\_primitive\_count}$$

$$\mathrm{Re}\,lative\_length\_error = 1 - P(R, \mu, \sigma)$$

Modified Probability Density Function:

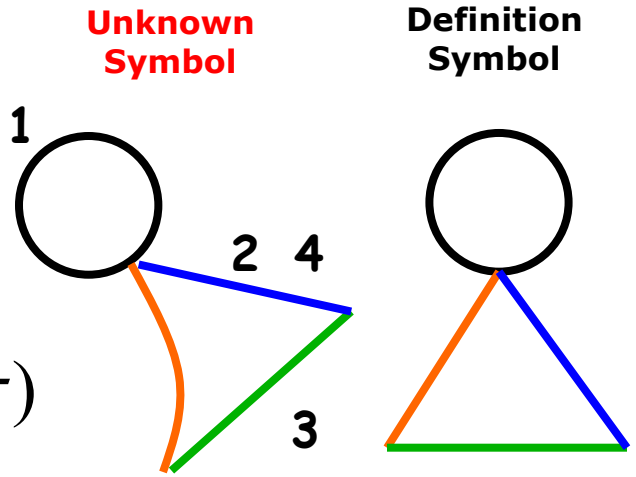
$$P(R, \mu, \sigma) = exp[-\frac{1}{50.0} \cdot \frac{(R-\mu)^4}{\sigma^4}]$$



Relative Length Error

$$E_3 = \frac{\sum \mathrm{Re}lative\_length\_errors}{MIN\_primitive\_count}$$

$$\mathrm{Re}lative\_length\_error = 1 - P(R, \mu, \sigma)$$



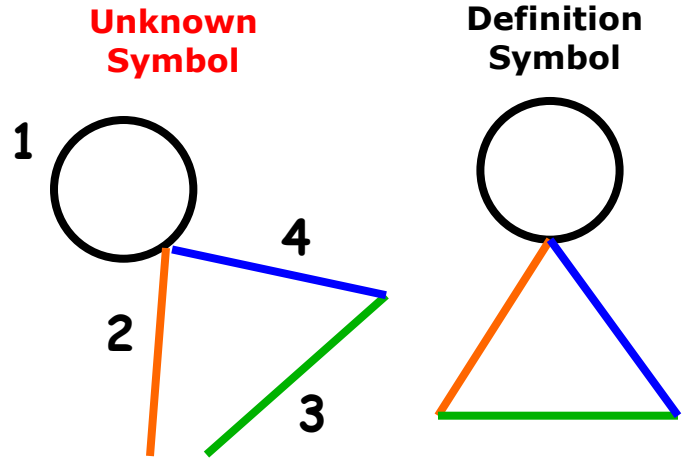
E_3 = ((1 - 1)
+ (1 - 0.73)
+ (1 - 0.73)
+ (1 - 0.98)) / 4
= 0.14

Node	U_R	μ	σ	$P(U_R)$
1	0.38	0.40	0.04	1.00
2	0.25	0.20	0.05	0.98
3	0.19	0.20	0.005	0.73
4	0.18	0.20	0.01	0.73

Number of Intersections Error

$$E_4' = \frac{\sum Intersection\_\#\_differences}{\#\_of\_pairs\_of\_primitives}$$

$$E_4 = Squash(E_4')$$



Example:

E'_4 =1/6
= 0.17

E_4 = S(0.17)
= 0.0068

	$I(U_i,U_j)$	$I(D_i,D_j)$
1-2	1	1
1-3	0	0
1-4	1	1
2-3	0	1
2-4	1	1
3-4	1	1

Intersection Angle Error

$$E_5 = \frac{\sum Errors\_in\_angles}{\#\_of\_angles}$$

Intersection Location Error

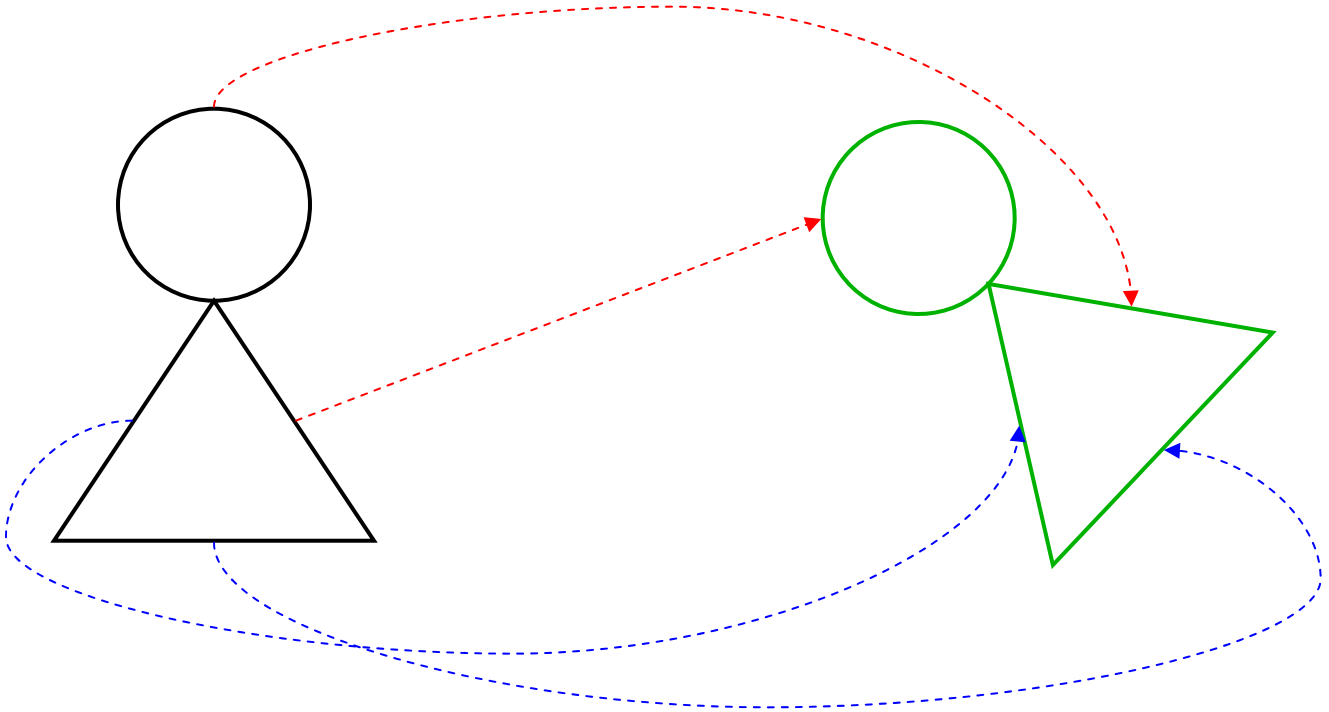
$$E_6 = \frac{\sum Errors\_in\_the\_intersection\_locations}{2 \times Number\_of\_intersections}$$

Roadmap

- ✓ Representation
- ✓ Measuring Similarity
- Graph matching
- User Study
- Conclusions

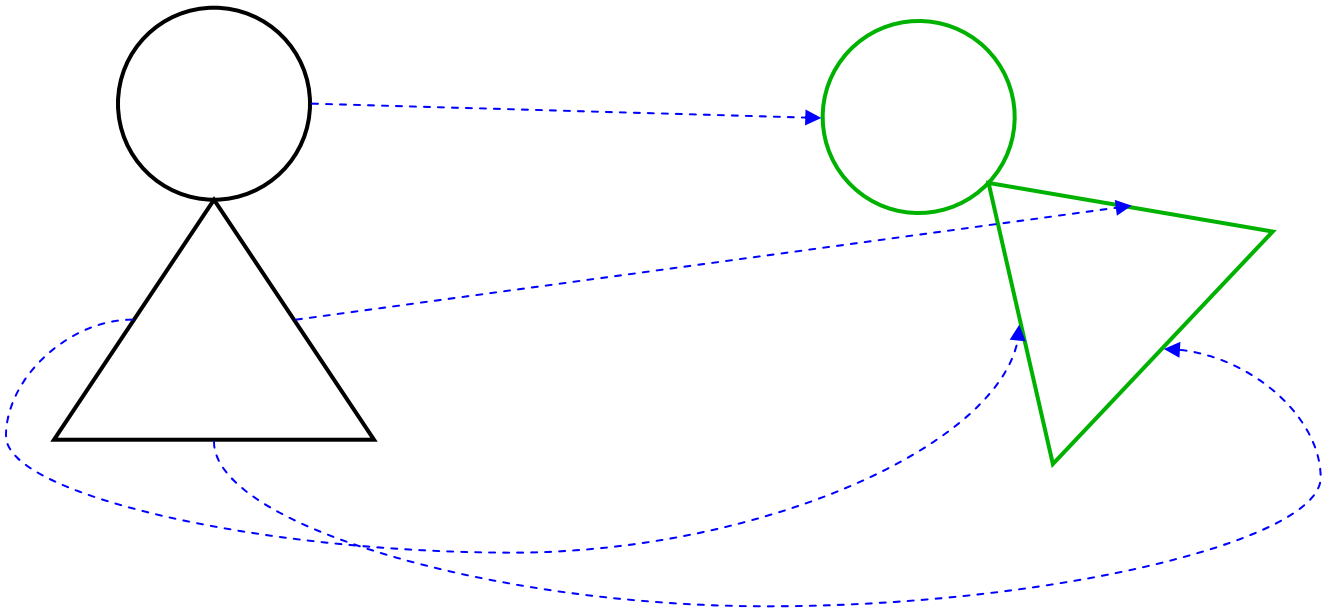
Graph Matching

Task: Find best node-pair assignment



Graph Matching

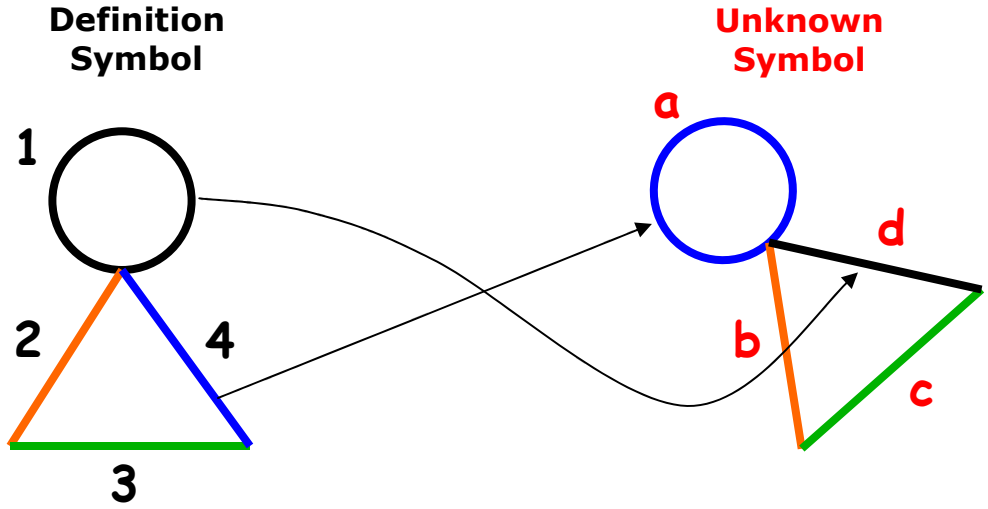
Task: Find best node-pair assignment



Graph Matching

- Approach: 4 approximate algorithms
  - Stochastic Matching
  - Error driven Matching
  - Greedy Matching
  - Sort Matching

Stochastic Matching

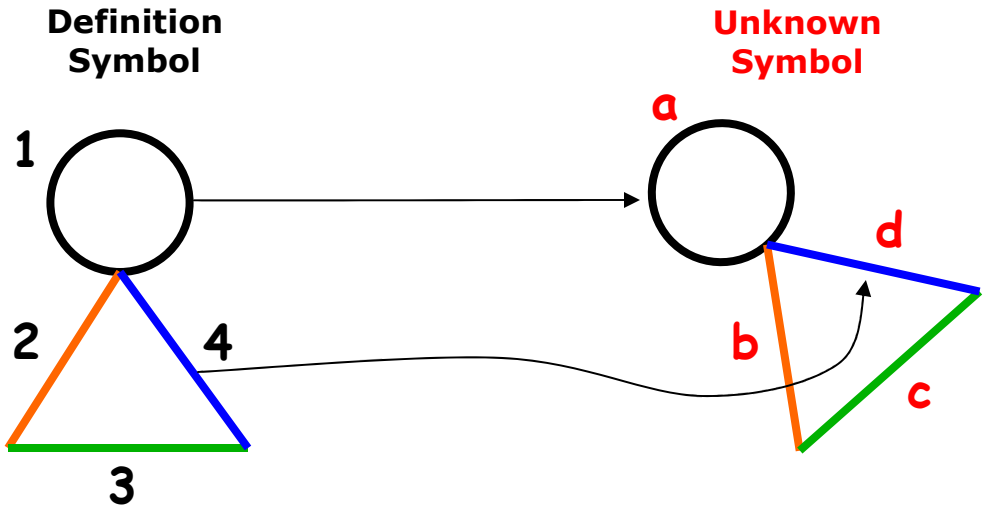


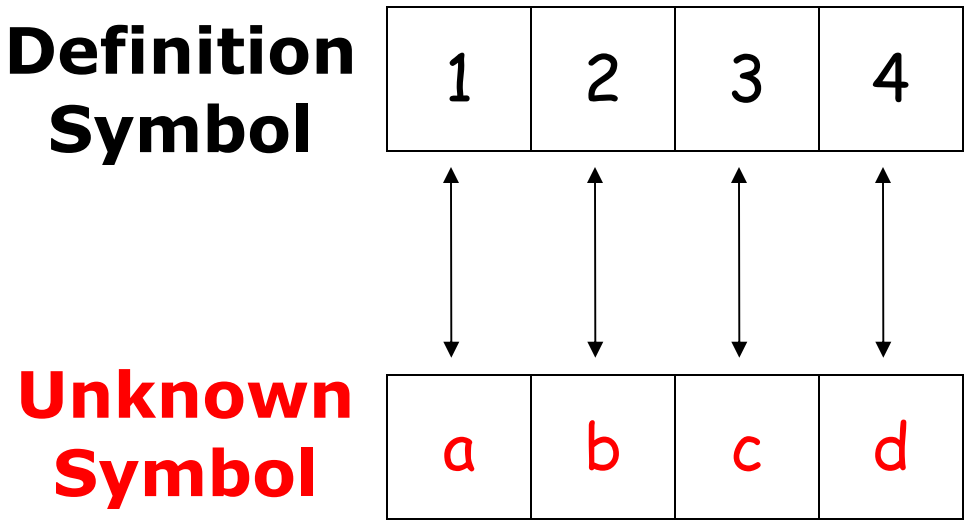
Definition Symbol

Unknown Symbol

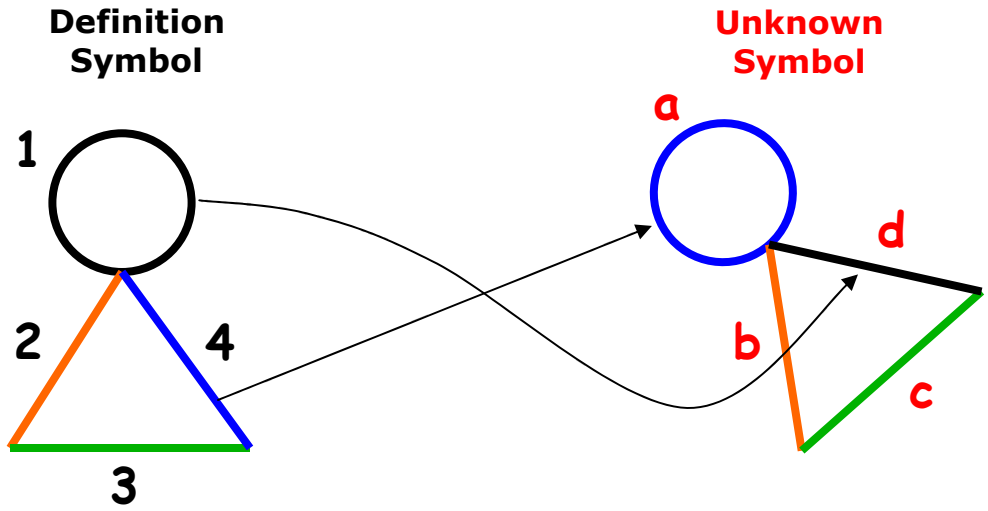
1	2	3	4
d	b	c	a

Stochastic Matching





Error-driven Matching

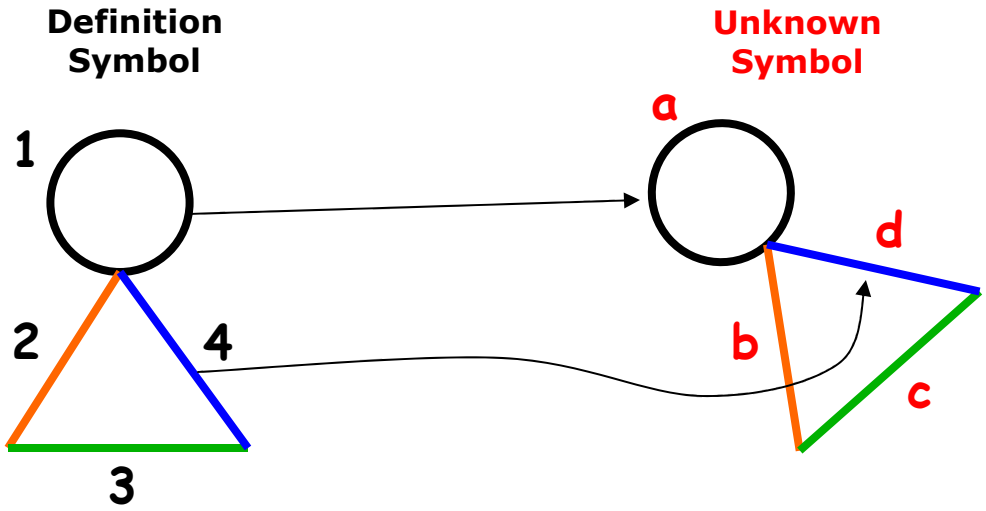


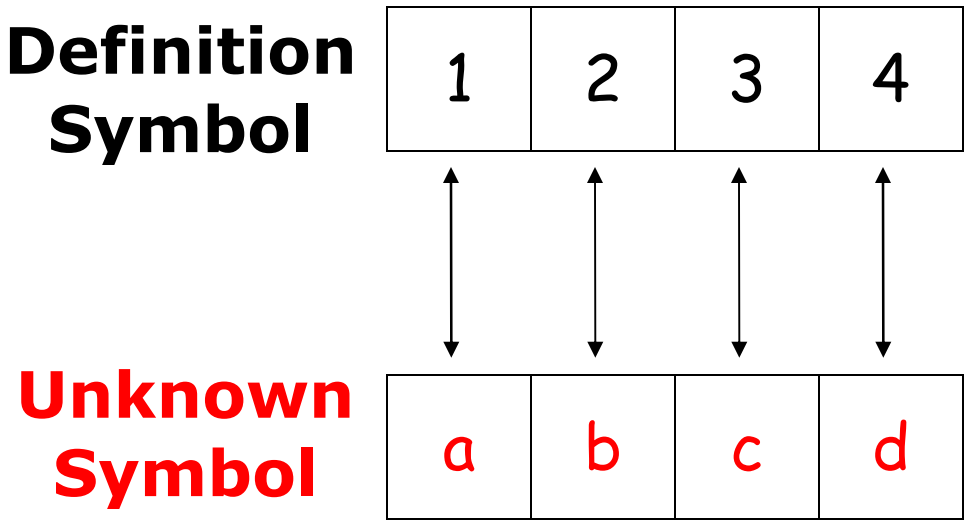
Definition Symbol

Unknown Symbol

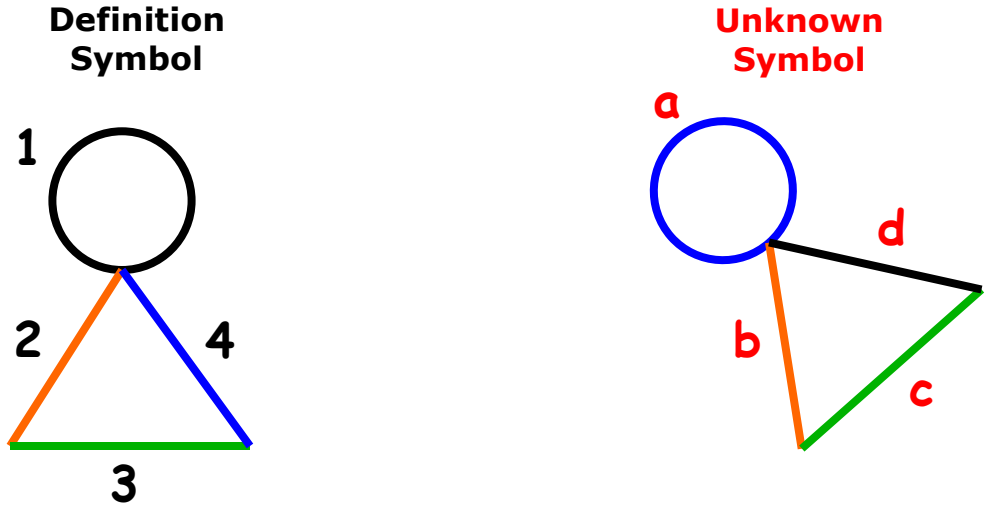
1	2	3	4
↕	↕	↕	↕
d	b	c	a

Error-driven Matching



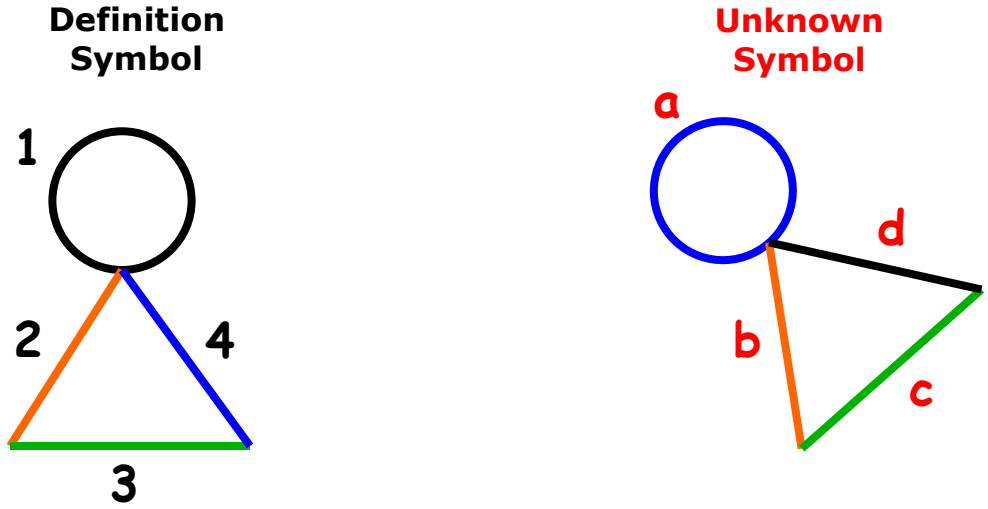


Greedy Matching



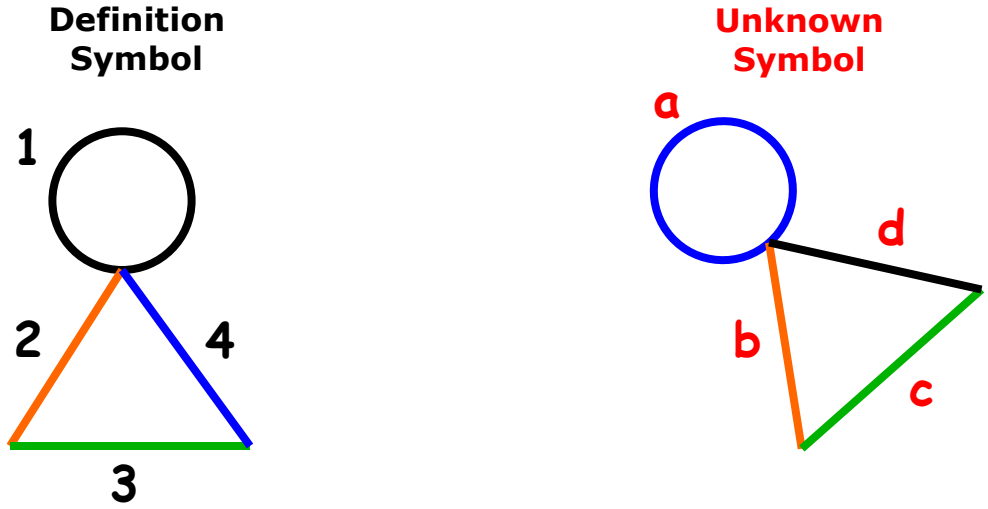
Definition Symbol	1	2	3	4
Unknown Symbol	d	b	c	a

Greedy Matching



Definition Symbol	1	2	3	4
Unknown Symbol	d	b	c	a

Greedy Matching

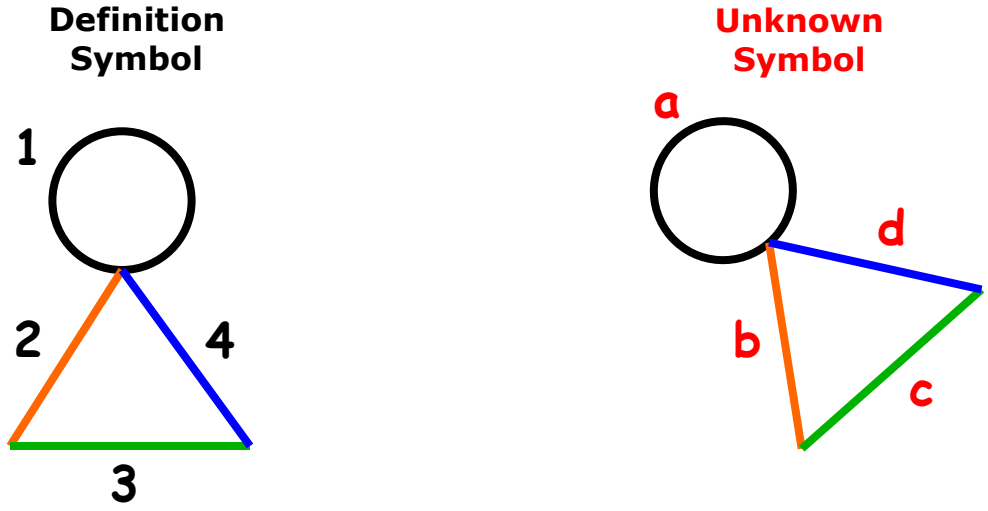


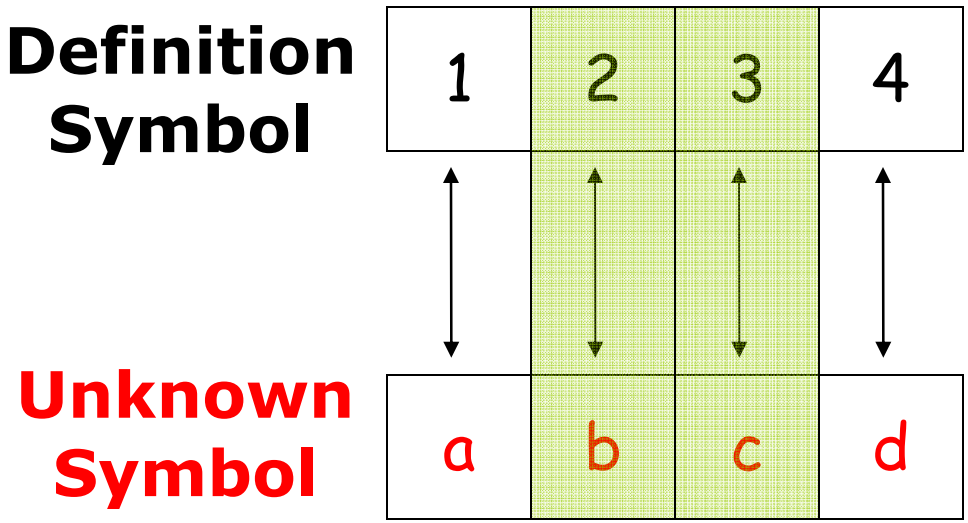
Definition Symbol

Unknown Symbol

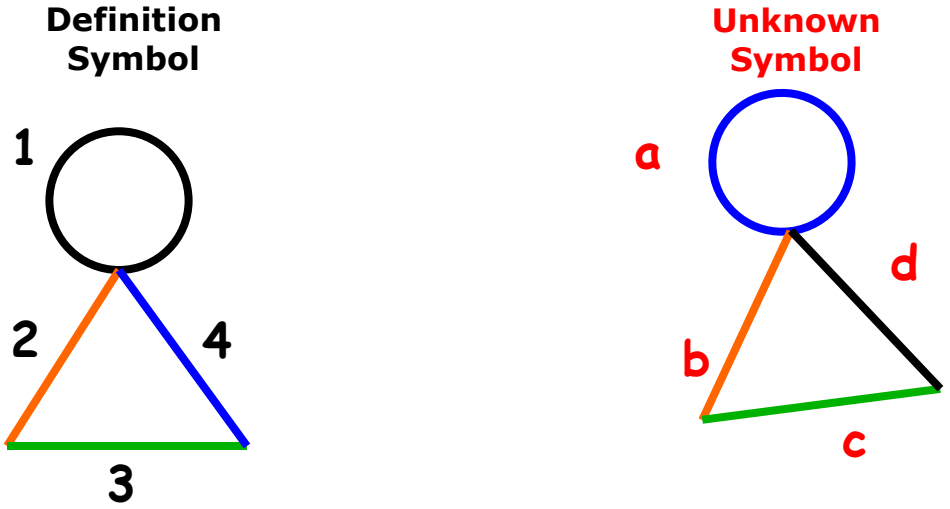
1	2	3	4
d	b	c	a

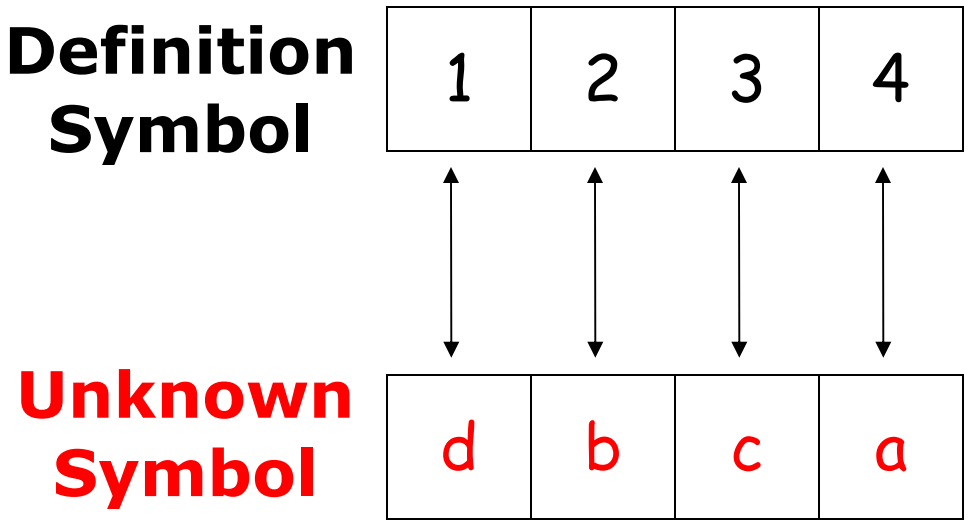
Greedy Matching



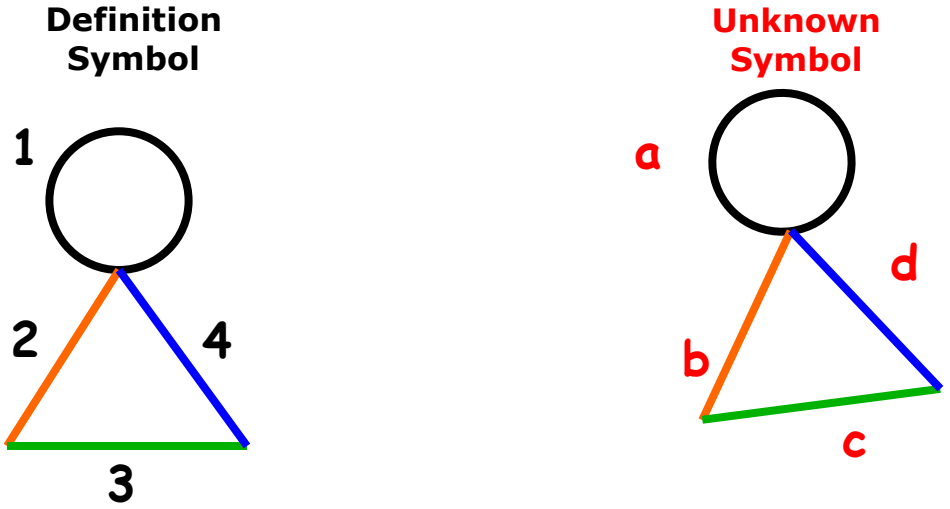


Sort Matching





Sort Matching



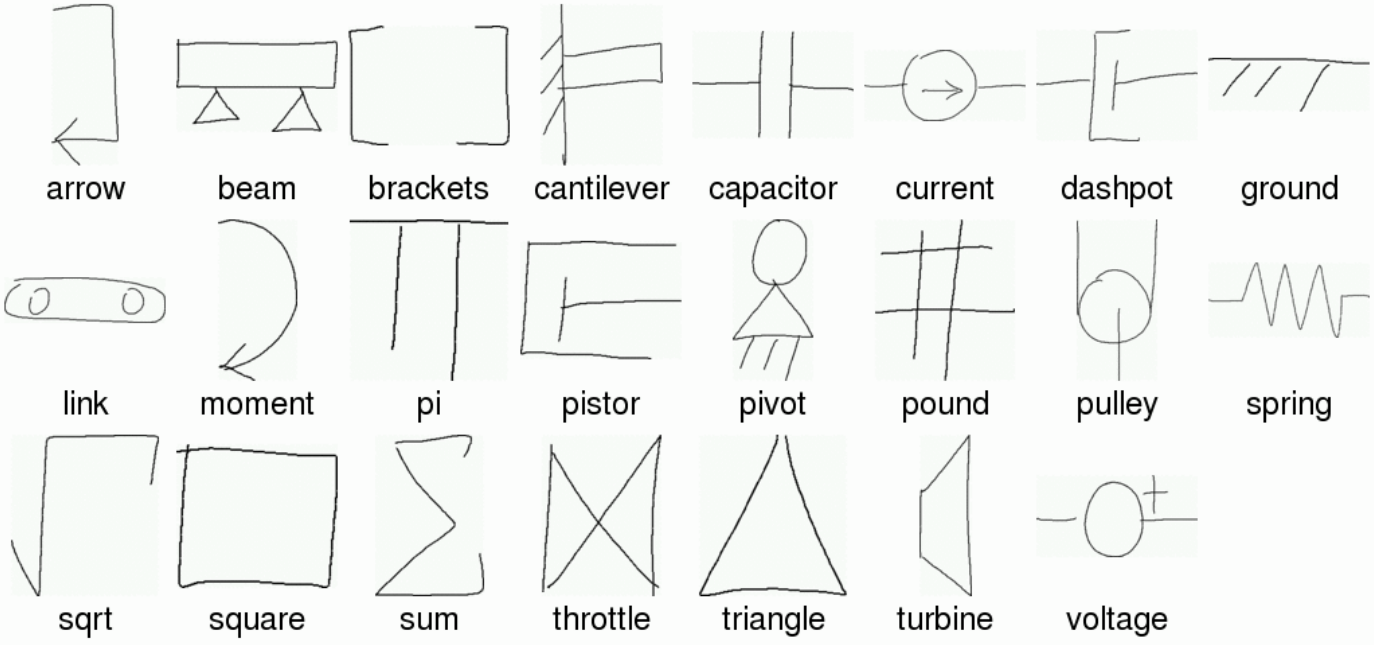
Definition Symbol	3	2	1	4
Unknown Symbol	c	b	a	d

Roadmap

- ✓ Representation
- ✓ Measuring Similarity
- ✓ Graph matching
- User Study
- Conclusions

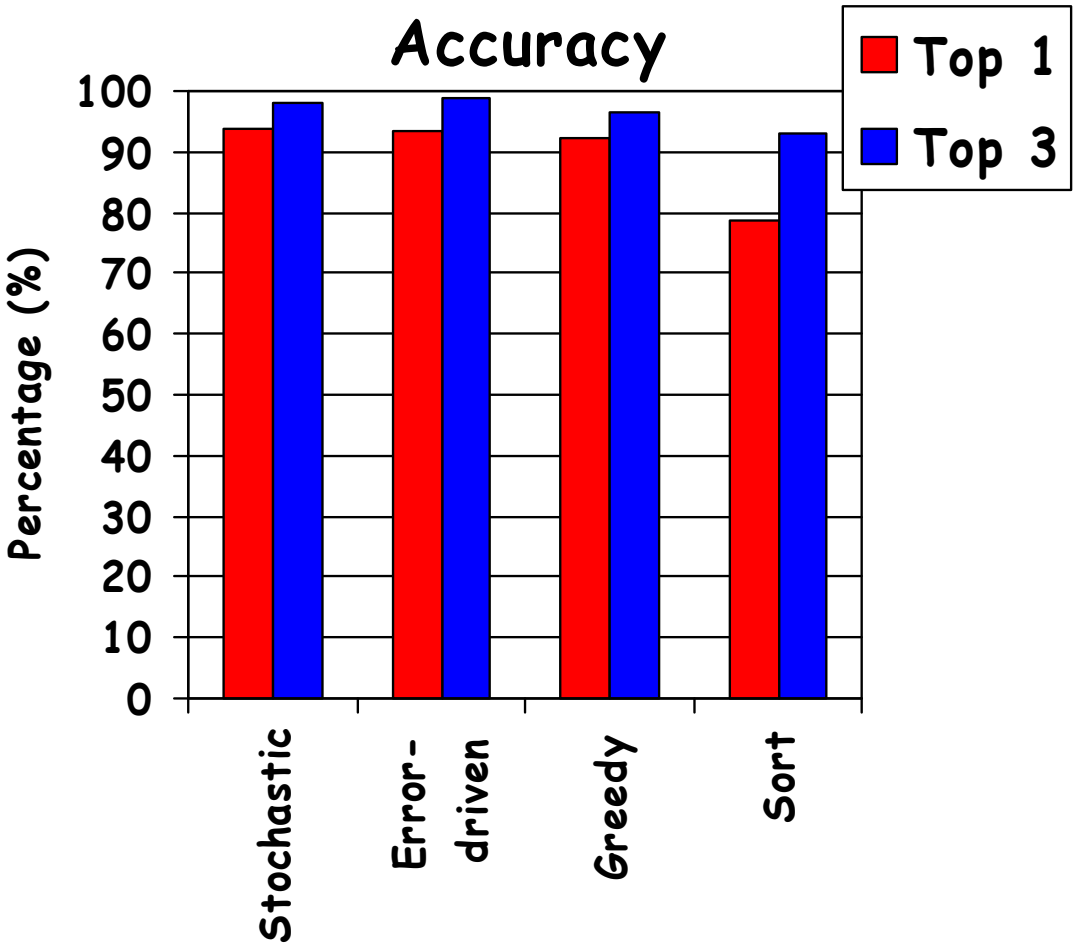
User Study

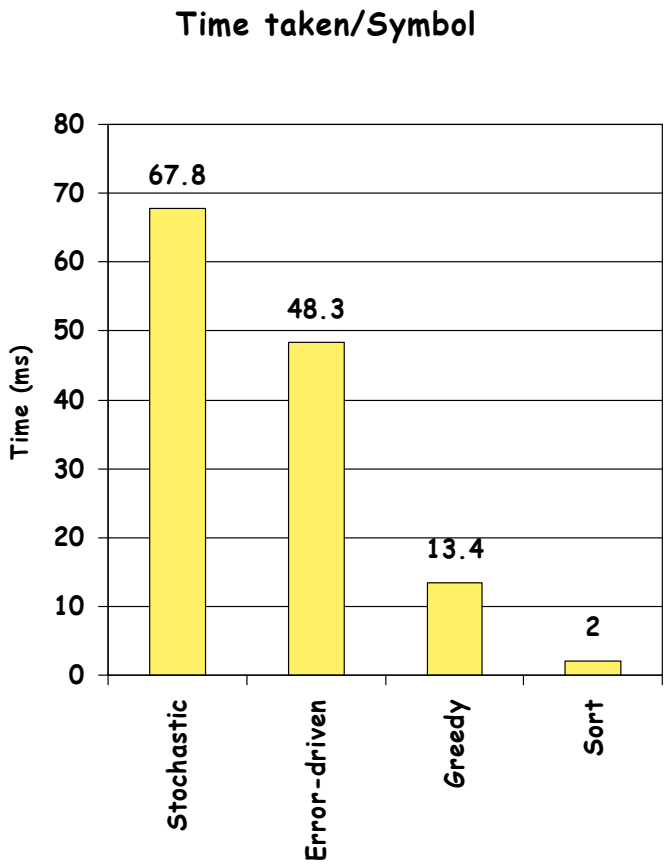
- 23 classes of mechanical symbols
- 9 participants
- 15 examples per symbol



Results

- On a P4 3.2G & 1G RAM machine





Conclusion

- Developed graph-based symbol recognizer
  - Insensitive to scale, orientation
- Stochastic Matching
  - Most accurate, most expensive
- Greedy Matching
  - Good compromise
- Sort Matching
  - Very fast, good for low-power devices

Q&A

Thank You!