When is a Problem Solved?

Daniel Lopresti

Computer Science & Engineering
Lehigh University
Bethlehem, PA, USA

George Nagy

Electrical, Computer, and
Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY, USA

If you have a QRcode reader on your smartphone, get it ready for later.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 1

LEHIGH
U N I V E R S I T Y

ICDAR 2011
Sept. 19, 2011

What's the problem?

We define our open problems as automating a task: this is quite different from math, physics, theoretical CS, etc.

Some ways of measuring success:
- Relative accuracy of new algorithm vs. previous methods.
- Relative accuracy of algorithm vs. expert human.
- Distinguishability of algorithm from <u>some</u> human.
- Current degree of community interest (publishability).
- Economic considerations (net payoff for using method).

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 2

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

What's the problem?

When is a problem solved?

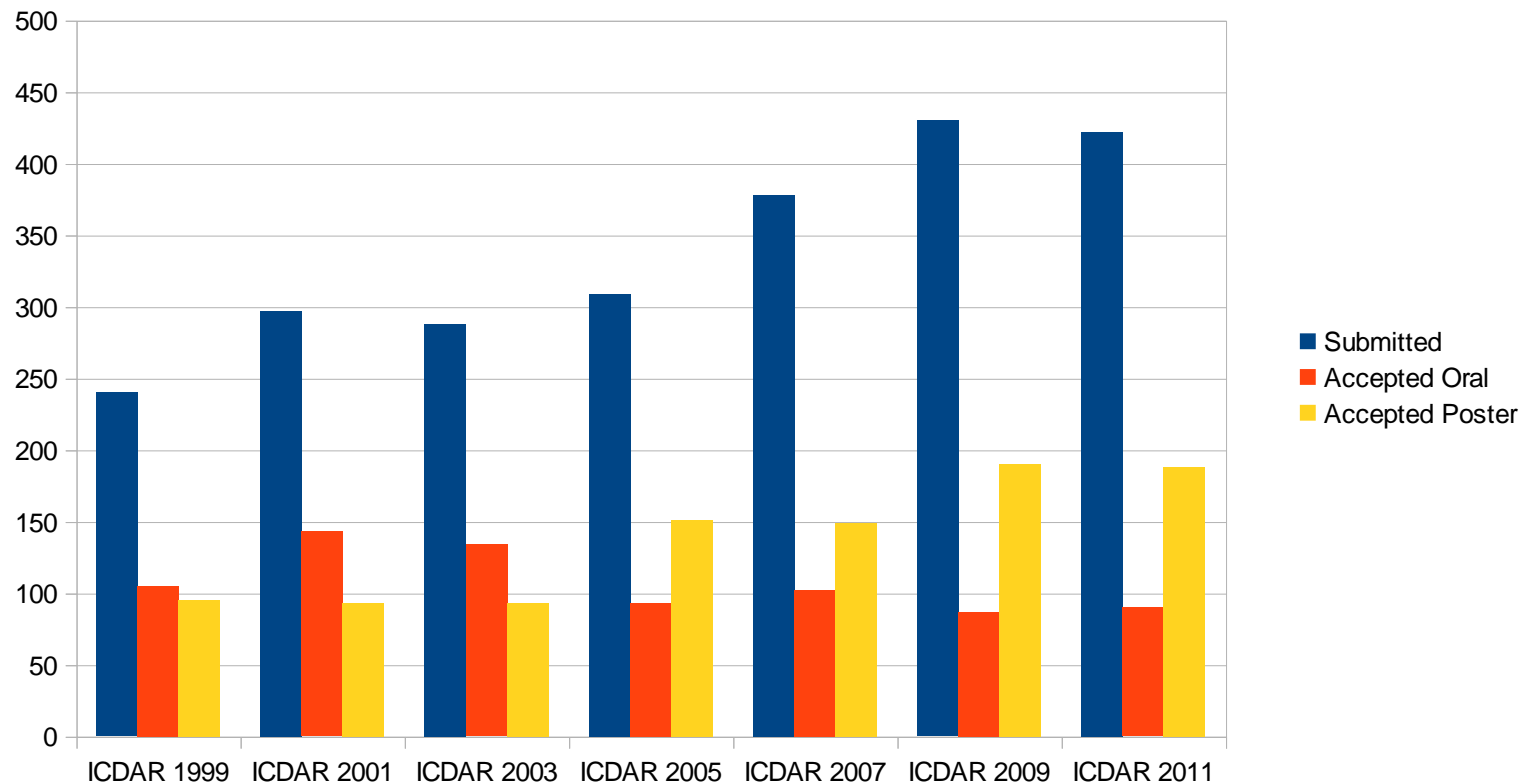This seems like a simple, basic question.

It also seems like an important question.

But we're not sure we know how to answer it ...

When is a Problem Solved?
Lopresti and Nagy
Rensselaer
Slide 3
LEHIGH UNIVERSITY™
ICDAR 2011
Sept. 19, 2011

Why should we care?

Not knowing the answer hasn't slowed us down:

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 4

LEHIGH
U N I V E R S I T Y ™

ICDAR 2011
Sept. 19, 2011

Why should we care?

Still, it should be troubling if we can't agree on an answer:

- wasted effort,
- missed opportunities,
- reduced relevance.

The answer seems to demand a community consensus.

Can there be a community consensus?

At the very least, it seems worthy of discussion.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 5

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Viewpoint #1

The endless pursuit of perfection:

"A problem is solved if there is a method which has been widely publicized and documented and freely available to the community which achieves 100% accuracy on within-spec inputs it receives."

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 6

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Viewpoint #2

The Turing Test:

> "A problem is solved if there is a method which has been widely publicized and documented and freely available to the community which generates output for a given input that a human judge cannot reliably distinguish from the output of a human expert."

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 7

LEHIGH
U N I V E R S I T Y

ICDAR 2011
Sept. 19, 2011

Viewpoint #3

As good as it gets:

"A problem is solved if there is a method which has been widely publicized and documented and freely available to the community which performs better than any other method, and which cannot be further improved without investing excessive resources."

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 8

LEHIGH
U N I V E R S I T Y™

ICDAR 2011
Sept. 19, 2011

Viewpoint #4

Good enough to get the job done:

> "A problem is solved if there is a method which has been widely publicized and documented and freely available to the community which cannot be replaced with any other method to improve the end-to-end performance of a specific application of interest."

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 9

LEHIGH
UNIVERSITY.

ICDAR 2011
Sept. 19, 2011

Viewpoint #5

Pure pragmatism:

> "A problem is solved when it is no longer possible to get a paper published on the topic (or, alternatively, to raise research funding to study the question)."

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 10

LEHIGH
U N I V E R S I T Y

ICDAR 2011
Sept. 19, 2011

Show of Hands

Which viewpoint(s) do you agree with?

- The endless pursuit of perfection.
- The Turing Test.
- As good as it gets.
- Good enough to get the job done.
- Pure pragmatism.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 11

LEHIGH
UNIVERSITY™

ICDAR 2011
Sept. 19, 2011

Key Considerations

When attempting to define problems and solutions, we must keep in mind the following important points:

- Populations and samples: performance figures like error, reject, or retrieval rates are of interest only with regard to populations rather than particular samples.

- Algorithms, heuristics, and implementations: most document analysis is built on heuristics rather than algorithms, although the latter term is applied to both. To be a solution, an algorithm must be implementable.

- Desirable criteria: solutions should be invariant to 90º rotation, modest differences in resolution, remapping RGB/gray values, jitters in threshold settings, etc.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 12

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Solved Problems?

At the risk of enflaming discussing (which is, in fact, our goal), we raise the following as potential candidates:

Yes	Nearly	No

Binarization.

Skew estimation.

Page segmentation into ≤5 types.

Thinning and skeletonization.

Printed paragraph-, line-, word-finding.

Handwritten paragraph-, line-, word-finding.

Context-free isolated character recognition.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 13

LEHIGH
U N I V E R S I T Y

ICDAR 2011
Sept. 19, 2011

Forcing the Question

Beyond just talk, what might result in decisive steps?

- Standardized datasets. Arguably these introduce as many issues as they seem to address.

- Competitions. Probably a good idea, although we have yet to see a problem declared solved as a result of one.

- Publishing policies. Journals and conferences could be more demanding when it comes to documentation, data-sharing, reproducibility, and comparison to the prior art.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 14

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Raising More Questions

How do we write about our progress? Simple analysis of accuracy claim statements for ICDAR'99 vs. ICDAR'09:
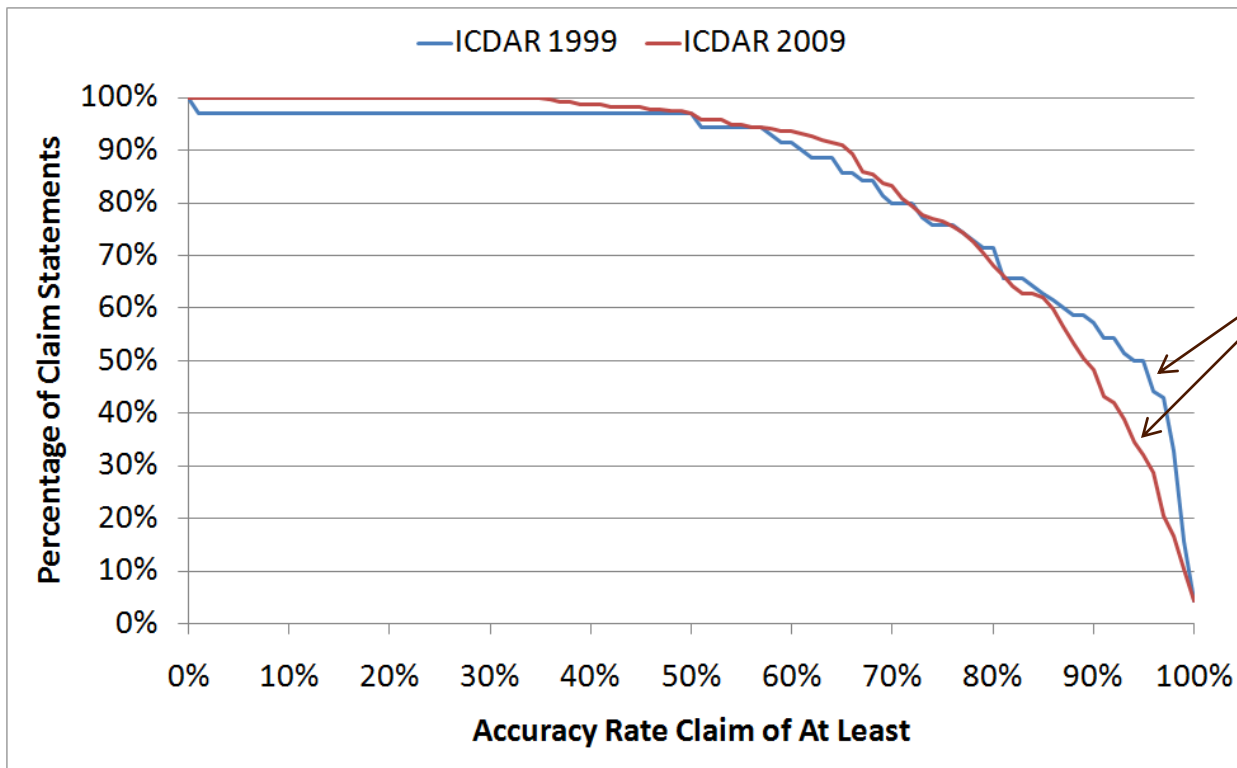
Are we getting more modest? Or working on harder problems?

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 15

LEHIGH UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Conclusions

Time seems ripe for a community discussion:

- When is a problem solved?
- Once we have an accepted definition, how to employ it?
- How do we begin to generalize our methods for applications beyond document image analysis?
- Beyond automating specific tasks, what else are we hoping to achieve?

What do you think?

Which problems do you consider solved?

Use the QRcode to participate in our anonymous, completely unscientific survey. The community's responses will be posted along with these slides one week after the end of ICDAR on DPL's website.

http://www.surveymonkey.com/s/H3XHKPT

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 17

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 1

1. What are appropriate measures of success?						Create Chart Download
	Strongly Agree	Somewhat Agree	Agree	Disagree	Strongly Disagree	Response Count
Net improvement of new algorithm over previous methods.	77.8% (7)	0.0% (0)	11.1% (1)	11.1% (1)	0.0% (0)	9
Relative accuracy of algorithm vs. human.	33.3% (3)	55.6% (5)	0.0% (0)	11.1% (1)	0.0% (0)	9
Increasing generality without lowering accuracy.	33.3% (3)	55.6% (5)	0.0% (0)	11.1% (1)	0.0% (0)	9
Current degree of community interest (publishability).	0.0% (0)	11.1% (1)	33.3% (3)	55.6% (5)	0.0% (0)	9
Economic considerations (net payoff for using method).	0.0% (0)	37.5% (3)	25.0% (2)	37.5% (3)	0.0% (0)	8
					answered question	9
					skipped question	0

Note: all results are informal and should be viewed as much for entertainment purposes as for enlightenment!

Results of Community Survey 1

What are appropriate measures of success?

Observations:

- "Net improvement" is a well accepted success measure.
- "Publishability" is rejected as an appropriate measure.
- Divergent opinions about "economic considerations."
- Somewhat less consensus about other measures.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 19

LEHIGH
U N I V E R S I T Y™

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 2

2. When should we consider a problem solved?				Create Chart Download
	Definitely solved	Might be solved	Not necessarily solved	Response Count
When there is a method that always achieves 100% accuracy on any input within specifications.	66.7% (6)	11.1% (1)	22.2% (2)	9
When there is a method that achieves 100% accuracy on well-behaved inputs.	33.3% (3)	44.4% (4)	22.2% (2)	9
When accuracy rates are high enough that any additional improvement is insignificant with respect to the targeted application.	77.8% (7)	11.1% (1)	11.1% (1)	9
When it is easy to obtain a reliable implementation to use in a general-purpose document analysis pipeline.	33.3% (3)	66.7% (6)	0.0% (0)	9
When there is a method that produces output that is indistinguishable from some human (not necessarily an expert).	11.1% (1)	55.6% (5)	33.3% (3)	9
When the community stops accepting papers on the topic in its conferences and journals.	0.0% (0)	22.2% (2)	77.8% (7)	9
			answered question	9
			skipped question	0

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 20

LEHIGH UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 2

When should we consider a problem solved?

Observations:

- Popular opinion: problem is solved when further improvement no longer helps intended application.
- 100% accuracy is another (obvious) criterion.
- Lack of publishability is not seen as a good measure.
- Less consensus about other possible criteria.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 21

LEHIGH UNIVERSITY™

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 3

3. What level of proof should be necessary before we consider a problem solved? 🥧 Create Chart ⬇ Download

	Necessary	Desirable	Optional	Unimportant	Response Count
Publication of a single compelling paper in a respected conference or journal.	44.4% (4)	11.1% (1)	44.4% (4)	0.0% (0)	9
Publication of a body of work by one researcher or team.	33.3% (3)	22.2% (2)	44.4% (4)	0.0% (0)	9
Extensive publication on the topic by multiple independent researchers over an extended period of time.	44.4% (4)	44.4% (4)	11.1% (1)	0.0% (0)	9
Availability of an open source solution implementing a well-studied algorithm.	22.2% (2)	77.8% (7)	0.0% (0)	0.0% (0)	9
				answered question	9
				skipped question	0

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 22

LEHIGH
UNIVERSITY

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 3

What level of proof is necessary … ?

Observations:

- Few criteria considered "necessary" – largest emphasis is on less demanding "desirable."

- Strongest opinions seem focused on open publication of a method, and open availability of an implementation.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 23

LEHIGH
U N I V E R S I T Y ™

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 4

4. Which of the following document image analysis problems do you consider "solved"? ⬤ Create Chart ⬇ Download

	Completely Solved	Mostly Solved	Not Solved	Not Sure	Response Count
Binarization.	33.3% (3)	55.6% (5)	0.0% (0)	11.1% (1)	9
Skew estimation.	33.3% (3)	33.3% (3)	11.1% (1)	22.2% (2)	9
Page segmentation into a small number of component types.	22.2% (2)	55.6% (5)	0.0% (0)	22.2% (2)	9
Thinning and skeletonization.	88.9% (8)	0.0% (0)	0.0% (0)	11.1% (1)	9
Machine-print paragraph-, line-, and word-finding.	55.6% (5)	11.1% (1)	11.1% (1)	22.2% (2)	9
Handwritten paragraph-, line-, and word-finding.	11.1% (1)	22.2% (2)	33.3% (3)	33.3% (3)	9
Context-free isolated character recognition.	44.4% (4)	44.4% (4)	0.0% (0)	11.1% (1)	9
				answered question	9
				skipped question	0

Results of Community Survey 4

Which of the following problems are solved?

Observations:

- Except for thinning and skeletonization (considered solved), little strong consensus.

- Problems involving handwriting (paragraph-, line-, and word-finding) seem to be considered unsolved.

- Some problems considered "mostly solved" – perhaps this is because we can always create hard cases.

When is a Problem Solved?
Lopresti and Nagy

Rensselaer

Slide 25

LEHIGH
UNIVERSITY™

ICDAR 2011
Sept. 19, 2011

Results of Community Survey 5

5. If you have other thoughts / comments / suggestions, please feel free to add them here. Thank you!   ⬇ Download

Response Count

Hide Responses   1

Responses (1)   Text Analysis   My Categories (0)

GOLD FEATURE: Text Analysis allows you to view frequently used words and phrases, categorize responses and turn open-ended text into data you can really use. To use Text Analysis, upgrade to a GOLD or PLATINUM plan.   [Learn More]   [Upgrade »]

Showing 1 text responses   No responses selected

We should probably combine several of the above metrics to determine when a problem is solved, e.g. I want an "all of the above" option.
10/31/2011 10:57 AM   View Responses

We consider this an excellent point, and thank all those who took the time to respond to our informal survey!