Can Natural Language Processing Become Natural Language Coaching?

Marti Hearst
UC Berkeley

ACL 2015 Keynote Presentation
Do you go to a yoga (or spin/karate/dance) class and just take notes?
Question: How best to memorize a text in 30 minutes?

What % time reading and studying? What % time practicing reciting from memory?

Answer: Study 1/3, Practice 2/3

“A curious peculiarity of our memory is that things are impressed better by **active** rather than by **passive** repetition.” -- William James, 1890
Twilight of the Lecture

The trend toward "active learning" may be ruled universities for 600 years.

by CRAIG LAMBERT

MARCH-APRIL 2012

EDUCATION

Farewell, Lecture?

Eric Mazur

Discussions of education are generally predicated on the assumption that we know what education is. I hope to convince you otherwise by recounting some of my own experiences. When I started teaching introductory physics to undergraduates at Harvard University, I never asked myself how I would educate my students. I did what my teachers had done—I lectured. I thought that was how one learns. Look around anywhere in the world and you’ll find lecture halls filled with students and, at the front, an instructor. This approach to education has not changed since before the Renaissance and the birth of scientific inquiry. Early in my career I received the first hints that something was wrong with teaching in this manner, but I had ignored it. Sometimes it’s hard to face reality.

When I started teaching, I prepared lecture notes and then taught from them. Because my lectures deviated from the textbook, I provided students with copies of these lecture notes. The infuriating result was that on my end-of-semester evaluations—which were quite good otherwise—a number of students complained that I was “lecturing straight from (his) lecture notes.” What was I supposed to do? Develop a set of lecture notes different from the ones I handed out? Ignore the students’ complaints?

A few years later, I discovered that students were right. My lectures, despite the high evaluations, were dull. The introductory physics course Newton presented. Every course can recite Newton's second law of motion, which states that the force of object A on object B in an interaction between two objects is equal in magnitude to the force of B on A—it sometimes is known as “action is reaction.” One day, when the course had progressed to more complicated material, I decided to test my students’ understanding of this concept not by doing traditional problems, but by asking them a set of basic conceptual questions (1, 2). One of the questions, for example, requires students to compare the forces that a heavy truck and a light car exert...
Active Learning Improves All Aspects of Learning

"Improved Learning in a Large Enrollment Physics Class",
Deslauriers et al., Science 2011

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Experimental Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance</td>
<td>53</td>
<td>75</td>
</tr>
<tr>
<td>Engagement</td>
<td>45</td>
<td>85</td>
</tr>
<tr>
<td>Learning</td>
<td>41</td>
<td>74</td>
</tr>
</tbody>
</table>
When to use Peer Discussion

1. Present content
2. Pose high-level, intellectual question
3. Students answer individually with clicker or flashcard
4. Analyze the distribution of responses
   - >70% correct: Provide brief explanation
   - 35%-70% Correct: Ask students to explain their reasoning to peer(s)
   - <35% Correct: Provide hint OR Provide detailed explanation
5. Students answer with clicker or flashcard
6. Provide explanation

Vickrey et al., CBE Life Sciences Education, 14, 2015
MOOCs (Massive Open Online Courses)
Emphasize Short Videos and Active Learning

Peter Norvig’s TED talk
250k certificates issued, 6.4 million course enrollments, 3 million users, 300 new courses offered this semester.

@edXOnline #FutureEDU

— Lorena Barba (@LorenaABarba) November 20, 2014
Why is Active Learning Like Coaching?

• Let students repeat problems until they get them right.
• Make failing fast, fun, easy to recover from.
• Provide as many practice problems as needed.
• Create dynamic difficulty levels that can be personalized based on how students are doing.
• Focus on the feedback rather than the grades.
Why Should NLP Researchers Care?

Hugely important problem

Intriguing new datasets

New approaches to algorithms
Some NLP in Education Organized Efforts

• Essay scoring and grading:
  • ACL/NAACL Workshops on Building Educational Applications Using NLP 2005-2015 (Tetreault, Burstein, Leacock)

• ESL correction:
  • CoNLL 2014 shared task (Ng et al.)
  • Helping Our Own Challenge 2012 (Dale and Kilgarriff)

• Special issue on resources and tools for language learners:
  • Sharoff et al. 2014, Language Resources and Evaluation 48(1).
Coaching for Writing Instruction

NLP for Peer Learning

Games for Learning Language

NLP for Automating Feedback
At the start of his book, I thought that the author did a pretty good job of setting up the scene but then again there were things that he also forgot to set up. I think that part of the problem was his failure to recognize the importance of the relationship between the characters. If I were to guess one thing about the story so far I would say that...
After I did some research, I knew that I would need some expert advice. Eventually, I went to a local dealership to check out some new models. I talked to the saleswoman and listened at she carefully. Her honesty and professionalism were really impressive. She had a lot of vary helpful suggestions and showed myself some safe affordable choices. After a long discussion I finally decided which one I wanted. She not only helped me with the paperwork and finished the sale, but also the insurance. I was expecting this purchase to be a serious hassle, but the experience was almost painless. Everything went smoothly, and now I have a brand new car!

I was so excited when I pulled out of the lot that the
the hike

She did the hike of a mountain. Then we went upstream to gorge. The beautiful day my boyfriend said. We went up to coffee.
WriteToLearn

LSA gives feedback as students revise essays.

3.5 revisions on average across ~21,000 students.

Student scores improve!

Suggests how to extend essay.

Foltz & Rosenstein, NIPS workshop, 2013
However, students don’t improve much on sentence fluency and conventions. Much of the improvement comes from doing the revisions.
WriteToLearn Video
Parsers are about Grammar \textit{Checking}

Checking for compliance with a grammar

In other words ... correcting.
How About Coaching?

• What would it take to make a parser that:
  • Reads alongside as the student writes, and
  • Tries to coach as choices were being considered, and
  • Considers common errors that might be about to happen, and
  • Learns the best way to suggest what explanations to show, and
  • Learns the best examples to show?

• What does grammar writing look like for this?
An Alternative: Real-Time Essay Coaching

• Imagine of thousands of students all writing the same essay
• Responding to coaching from program and teacher
An Alternative: Real-Time Essay Coaching

• The reactions of one set of writers acts as input for the next set

• If many students have a similar problem in a similar context:
  • Design an A/B test on the fly
  • See which suggestions help students out best
  • This will be better than “awk”!
Summary: Improving Writing Instruction

• New way to think about writing parsers
  • New grammars – one per grade level? One per L2L learning pair?
  • Integrate theories of sentence understanding into parsers?

• Opportunities for much of NLP
  • Paraphrasing
  • Rhetorical structure theory, etc

• Exciting opportunity for new kind of data collection
  • Hold the semantics constant, vary the expression
  • Observe the writing while it is happening, vary it in real time
Coaching for Writing Instruction

NLP for Peer Learning

NLP for Automating Feedback

Games for Learning Language
Tip Tap Tones:
Game to Learn Chinese Tones

Edge et al., MobiHCI 2012
chan vs zhan:
a green background means the user got the previous answer correct.
Red means incorrect.
4-way comparison
Tip Tap Tone Video
ToneWars
Game to Learn Chinese Tones *With Other People*

Non-native speakers get paired up with native speakers in real time.

Head et al., ITS 2014
ToneWars Video
Engkoo:
Synthesized Karaoke To Teach English

Wang et al, IEEE Computer 2012
Engkoo Video
Anand, Shreyas, Sharma, Starostenko, DeSouza, Ryokai, and Hearst, under review.

Enhance children’s metalinguistic awareness via visual feedback.
1. help old man
2. help Police man
3. help auto man
4. help bus man
5. help car man
Creating the App
Use reading wordlists

Visual words from a farm context

Subject - Verb - Object structure

Word cards to allow language manipulations
Visualizing Words

sheep = sheep_skin + sheep_mouth + sheep_eyes
Adding Emotions

sheep_skin_positive

sheep_mouth_happier

sheep_mouth_happy

sheep_mouth_sadder

sheep_mouth_sadder

sheep_mouth_angry

sheep_eyes_happier

sheep_eyes_happy

sheep_eyes_sadder

sheep_eyes_sadder

sheep_eyes_angry

sheep_eyes_asleep
Positioning the Object

- sky
- ground
- Vanishing Plane
  - back
  - middle
  - front
- left
- center
- right
Building Grammar Rules

A Cat is walking behind the fences

5 word piece
Building Grammar Rules

The Cats

A Cat

are

is

walking

behind

with

near

the fences
Building Grammar Rules
How do we make sentences come alive?
rolling
Horizontal Movement

start  mid  end
Horizontal Movement

start | mid | end
What did children understand?
Beginner (4 kids, aged 4 - 6 yrs)

Has trouble reading. Is still trying to understand how words form a sentence.

“P I G is pig”

“If there is cat, I should not use ‘are’”?

“in 3 words, you use 1 color for each blank”

Allows exploration to understand which letters correspond to which word
Helps learn new words
Helps understand how determiners modify nouns
Intermediate (7 kids, aged 5.5 - 8 yrs)

Can read simple words. Learning to make sentences.

“Amused means… super happy?”

“How can fierce be smiling?”

“Barns roll…. because it’s make believe”

Explained inconsistencies in images (which only showed one of two emotions). Made “silly” sentences and articulated why sentences were funny.
Expert (5 kids, aged 7.5 - 8 yrs)

Can read, and has worked on sentence diagramming.

“Depressed means sad”

“Purple words are the ones doing the action”

“When there is ‘with’, we have to use another animal. It cannot be fence”

Learned meanings of new words through the context.
Did not want to make “silly” (illogical) sentences.
Connected parts of speech to colors of cards.
Summary: Games to Learn Language

• Games engage kids (and adults!) to encourage practice

• If designed correctly, they can increase learning as well

• NLP research can take game-based learning far beyond vocabulary
Coaching for Writing Instruction

NLP for Peer Learning

Games for Learning Language

NLP for Automating Feedback
Mastery Learning
Mastery Learning Requires Feedback
NLP can help with identifying how many concepts students capture in exam essays.
Mastery Learning: Algebra Challenge

Liu, Ballweber, O’Rourke, Butler, Thummaphan, Popovic, TOCHI 2015.
Suzy is ten years older than Billy, and next year she will be twice as old as Billy. How old is Suzy now?

Evelyn went to the store 8 times last month. She buys 11 stickers each time she goes to the store. How many stickers did Evelyn buy last month?

You attended high school for 4 years. Each year you bought 7 new textbooks. How many textbooks do you have at home now?
Alternative: Make Word Problems Into Stories

• Let kids choose the characters and the themes
  • Fantasy, Science Fiction, Wizardry
• Use templates and constraint solvers to enforce logical relations
• Add in Discourse Tropes and Co-reference Resolution
Knight Alice has 30 chalices. Dragon Elliot has 9 chalices. Alice slays the dragon, and takes his chalices. How many chalices does she have now?
Summary: NLP for Feedback

• NLP can be used for mastery learning in at least 2 major ways:
  • Aiding in creating assessment rubrics
  • Automatically creating questions and answers for problems
Coaching for Writing Instruction

NLP for Peer Learning

NLP for Automating Feedback

Games for Learning Language
Learning With Other Students

Peers remember what it’s like to not understand

The way to really learn something is to teach it to someone else

Hundreds of research papers show the value of peer learning.
Peer Learning: Students talk with Peers
Discussion forum / Large lecture

Small group discussion
MOOCs from Students’ Perspectives

• Pros:
  • Free
  • Self-paced
  • Any time, any where

• Cons:
  • Isolating
  • Impersonal
  • No cohort to motivate
Our Experiments: Does Small Group Discussion Improve Student Experience in MOOCs?

• Hypothesis:
  • People working in groups will get the answers right more often than those working alone.
  • Especially if an incentive is given to work together.

• Expectations:
  • Discussions will be substantive (in-depth, on-topic)

Coetzee, D., Lim, S., Fox, A., Hartmann, B., and Hearst, M.A., CSCW 2015
Question

With the decline of predators, such as wolves and coyotes, that used to keep the deer population within certain limits, deer have increased in numbers until they cannot feed themselves in the forest alone but must forage on open rangeland in competition with cattle. Thus, in areas where forest borders on rangeland, deer hunting is an essential activity.

This argument would be most seriously weakened if it could be shown that

Choose one of A to E. Please scroll down if your screen does not display all choices.

Possible Answers

A  deer hunters are not concerned about the prosperity of ranchers
B  wolves and coyotes do not prey upon deer only
C  [Correct] deer and cattle do not eat the same plants
D  deer hunting is popular even in areas where the forest does not border rangeland
E  the deer population may someday be hunted out of existence

Discussion

Student 3: I chose C because it seemed to clash with the statement "deer hunting is an essential activity" more than other statements

Student 1: I don't think the popularity of deer hunting is the issue.

Student 3: Right, it's whether it's essential. If the deer and cattle don't eat the same plants, then deer hunting isn't essential to preserve the rangeland for the cattle.

Student 1: That's how I see it.

Me: Actually, I did not understand the question until now. Whoops.

Me: I actually do believe it's C.

Student 3: Awesome!

Me: That was easy!

Student 3: We all agree so let's hope for the bonus :)

Timer

03:10

Your first choice was D

Your final choice is C
MOOCChat Video
Results: Discussion Improves Results

- Higher % of correct final responses for workers in groups (Fisher’s test, p < 0.01)
- Bonus incentive also improved results
Results: Subjective impressions

• Most workers rated as enjoyable, and left positive feedback
• Similar results when deployed in real online course (53% rated enjoyable)
Sample Discussion

• Student 2: I think E is the right answer
• Student 1: Hi, I think E is right, too
• Student 3: Hi! This seems to be a nurture vs nature question.
• Student 3: Can scent be learned, or only at birth?
• Student 2: Yeah, but answer A supports the author's conclusion
• Student 1: I felt that about A too
• Student 2: But the question was, which statement would weaken the author's conclusion
• Student 3: So I choose A, showing that scent can be learned at not only AT BIRTH.
• Student 2: That's why I think E is right
• Student 3: Are you real, or fake?
• Student 2: real
• ... 
• Student 1: So, do we all agree that E was the correct answer?  
• Student 2: I think so, yes. 
• Student 3: But I'm sticking with A since "no other water could stimulate olfactory sites" and I suggest that other water could be detected. 
• Student 3: *and
• Student 1: I thought about c for awhile but it didn't really seem to have anything to do with the topic of scent 
• ... 
• Student 3: E is still about spawning ground water, I think. this is a confusing question.
• Student 1: I thought E contradicted the statement the most 
• Student 2: me too 
• Student 3: I loving hits with other mturkers
Lots of NLP Work To Do!

• Assess what level in the Bloom taxonomy the discussion is at

• Automatically detect levels of understanding
  • To assign TAs to discussions
  • To re-group students based on understanding
  • To create new quiz questions on the fly

• Discourse analysis to shape the discussion
Summary
Coaching for Writing Instruction

NLP for Peer Learning

NLP for Automating Feedback

Games for Learning Language
Conclusions: NLP to Coach Learning:

- Hugely important problem
- Intriguing new datasets
- New approaches to algorithms
Thank you!

Marti Hearst
UC Berkeley

Illustrations by www.justeleanor.com
Used with permission