

Making Reading More Effective

Technologies to Help Information Seekers

Sumit Basu, Lucy Vanderwende, Lee Becker*, Chuck Jacobs

Who Reads to Learn?



Mary
home buyer



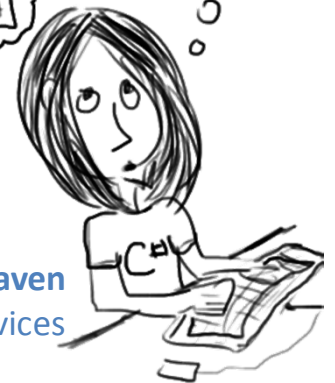
Jae
interview prep



Karthik
diabetes
management



Nina
grad student



Raven
web services

Common Threads:

- Self-motivated learners
- Wide variety of sources
- Factual and conceptual material
- A need for mastery

Making Reading More Effective

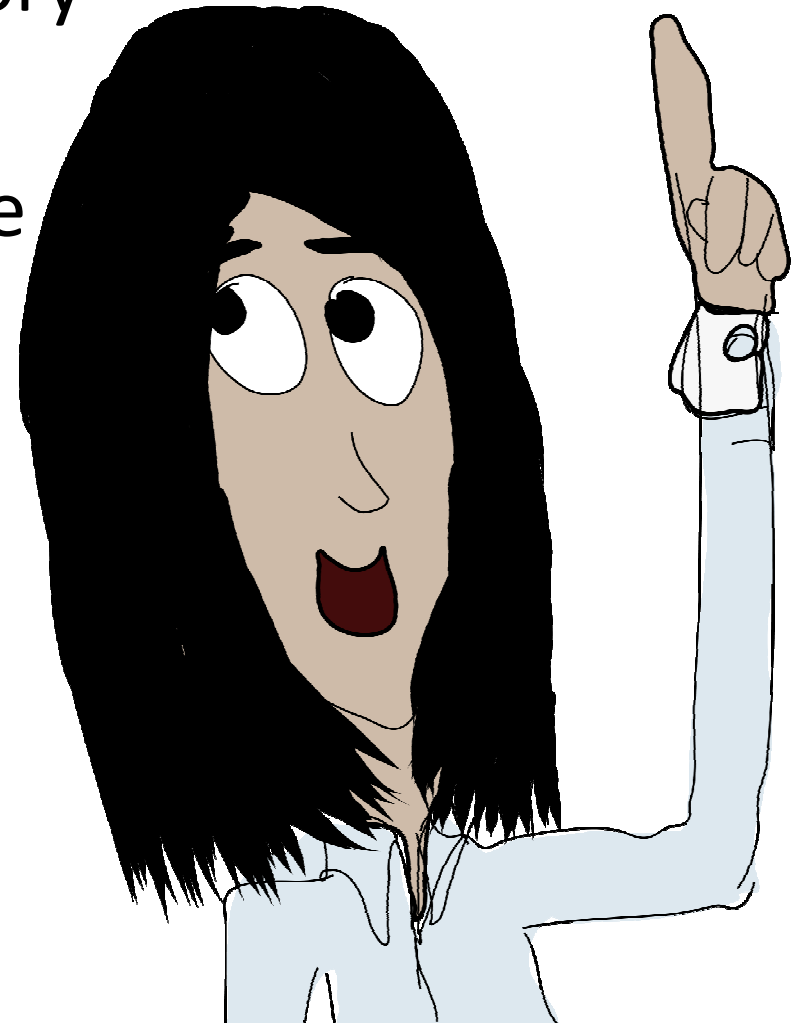
Axis 1: Improving Mastery

Axis 2: Improving Coverage

Axis 3: Improving Engagement

A Call for Collaborators and Interns

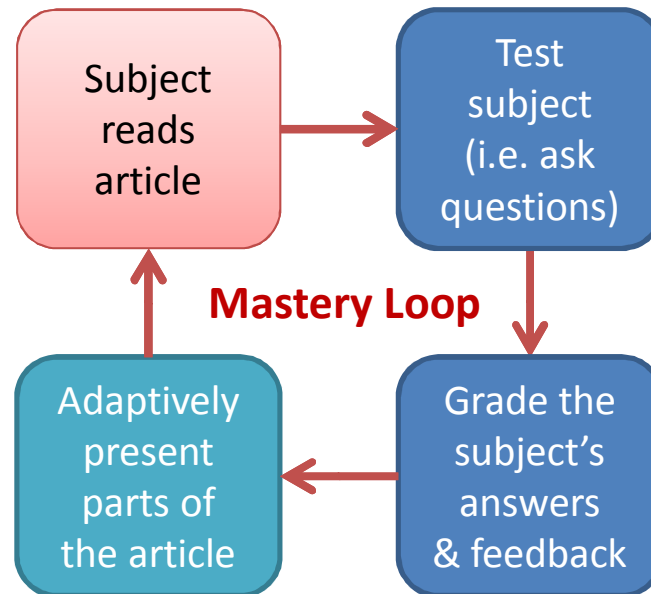
- A great deal of open territory
- If you see something that piques your interest, please contact us!
- sumitb@microsoft.com
lucyv@microsoft.com



Axis 1: Mastery

Methods for
Improving
Engagement

Question
Generation



Modeling of
Knowledge and
Coverage

Automatic
and Assisted
Grading

Mastery: The Value of Testing

The Critical Importance of Retrieval for Learning

Jeffrey D. Karpicke^{1*} and Henry L. Roediger III²

Learning is often considered complete when a student can produce the correct answer to a question. In our research, students in one condition learned foreign language vocabulary words in the standard paradigm of repeated study-test trials. In three other conditions, once a student had correctly produced the vocabulary item, it was repeatedly studied but dropped from further testing, repeatedly tested but dropped from further study, or dropped from both study and test. Repeated studying after learning had no effect on delayed recall, but repeated testing produced a large positive effect. In addition, students' predictions of their performance were uncorrelated with actual performance. The results demonstrate the critical role of retrieval practice in consolidating learning and show that even university students seem unaware of this fact.

Ever since the pioneering work of Ebbinghaus (1), scientists have generally studied human learning and memory by presenting people with information to be learned in a study period and testing them on it in a test period to see what they retained. When this procedure occurs over many trials, an exponential learning curve is produced. The standard assumption in nearly all research is that learning occurs while people study and encode material. Therefore, additional study should increase learning. Retrieving information on a test, however, is sometimes considered a relatively neutral event that measures the learning that occurred during study but does not by itself produce learning. Over the years, researchers have occasionally argued that learning can occur during testing (2–6). However, the assumptions that repeated studying promotes learning and that testing represents a neutral event that merely measures learning still permeate contemporary memory research as well as contemporary educational practice, where tests are also considered purely as assessments of knowledge.

Our goal in the present research was to examine these long-standing assumptions regard-

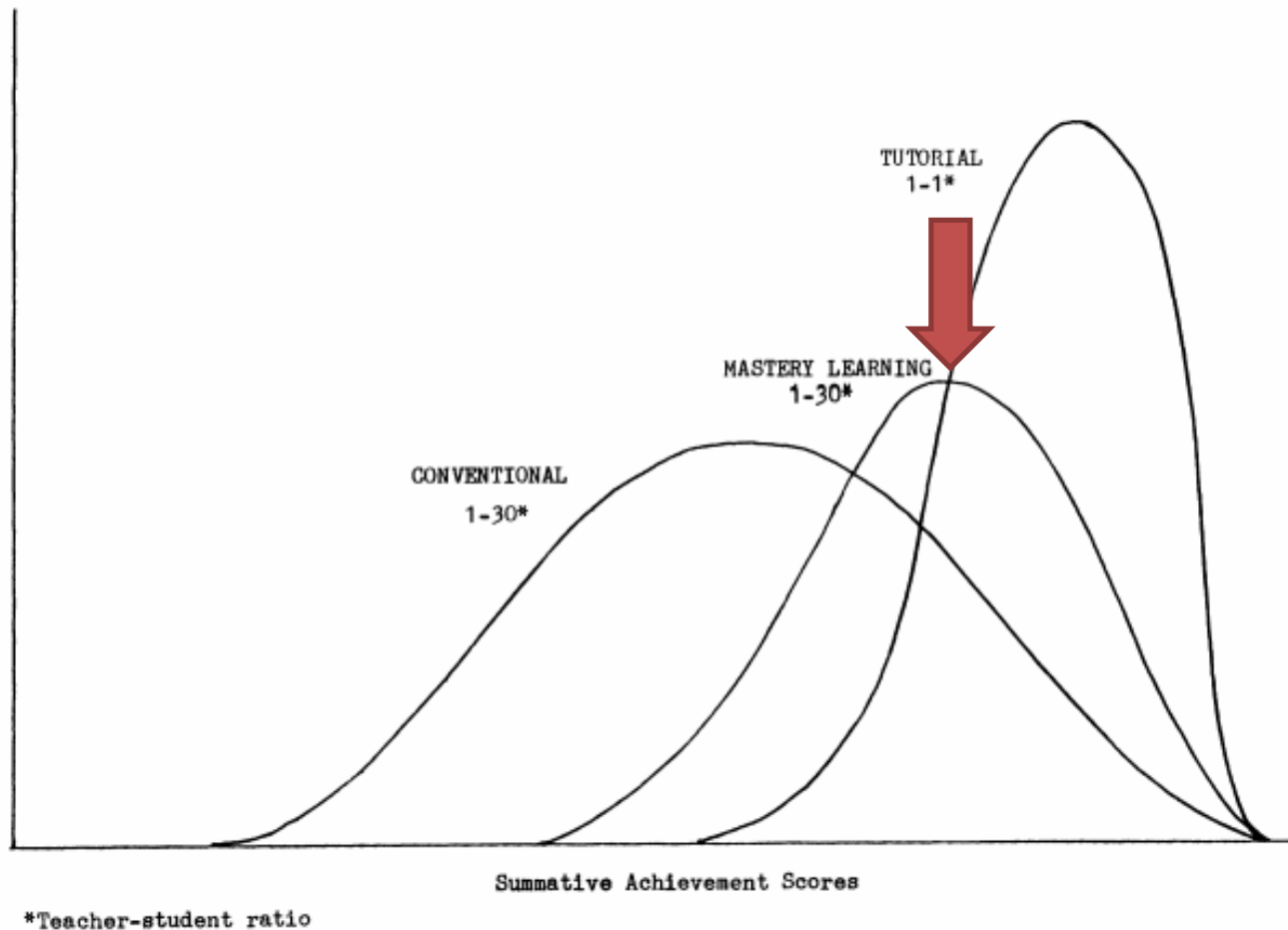
memory, concerning the relation between the speed with which something is learned and the rate at which it is forgotten. Is speed of learning correlated with long-term retention, and if so, is the correlation positive (processes that promote fast learning also slow forgetting and promote good retention) or negative (quick learning may be superficial and produce rapid forgetting)? Early research led to the conclusion that quick learning reduced the rate of forgetting and improved long-term retention (7), but later critics argued that, when forgetting is assessed more properly than in the early studies, no differences exist between forgetting rates for fast and slow learning conditions (8, 9). By any account, conditions that exhibit equivalent learning curves should produce equivalent retention after a delay (9).

Using foreign language vocabulary word pairs, we examined the contributions of repeated study and repeated testing to learning by comparing a standard learning condition to three dropout conditions. The standard method of measuring learning, used since Ebbinghaus's research (1), involves presenting subjects with information in a study

- **Karpicke and Roediger, 2008, “The Critical Importance of Retrieval for Learning.”**
- **Anderson and Biddle, 1975, “On Asking People Questions About What They are Reading.”**
- **Laufer and Goldstein, 2004, on the difficulty of Recall tasks vs. Recognition**
- **The Dunning-Kruger effect: the cognitive bias in which the unskilled think they have mastery**
- **McGraw-Hill representatives – the persistent need for new tests for teachers (helper tool) and students (self-review)**

Mastery: The Value of Adaptation

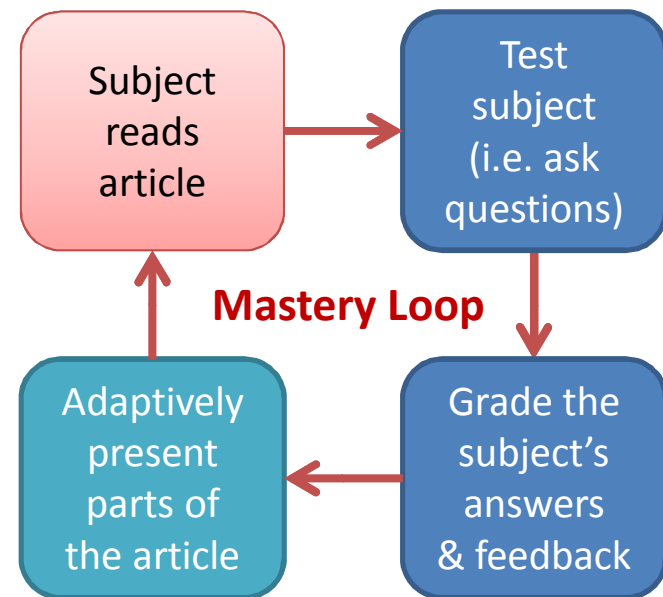
FIGURE 1. Achievement distribution for students under conventional, mastery learning, and tutorial instruction.



From Bloom et al., "The Two Sigma Problem," *Educational Researcher*, 13 (6) pp. 4-16. 1984.

So, What Can We Do?

- Mastery is achieved through a repeated cycle of testing and adaptive presentation
- Our work is focused on making it possible to apply the Mastery Loop at scale via:
 - Automatic methods
 - Auto/Crowdsourcing hybrids
 - Amplifying human efforts



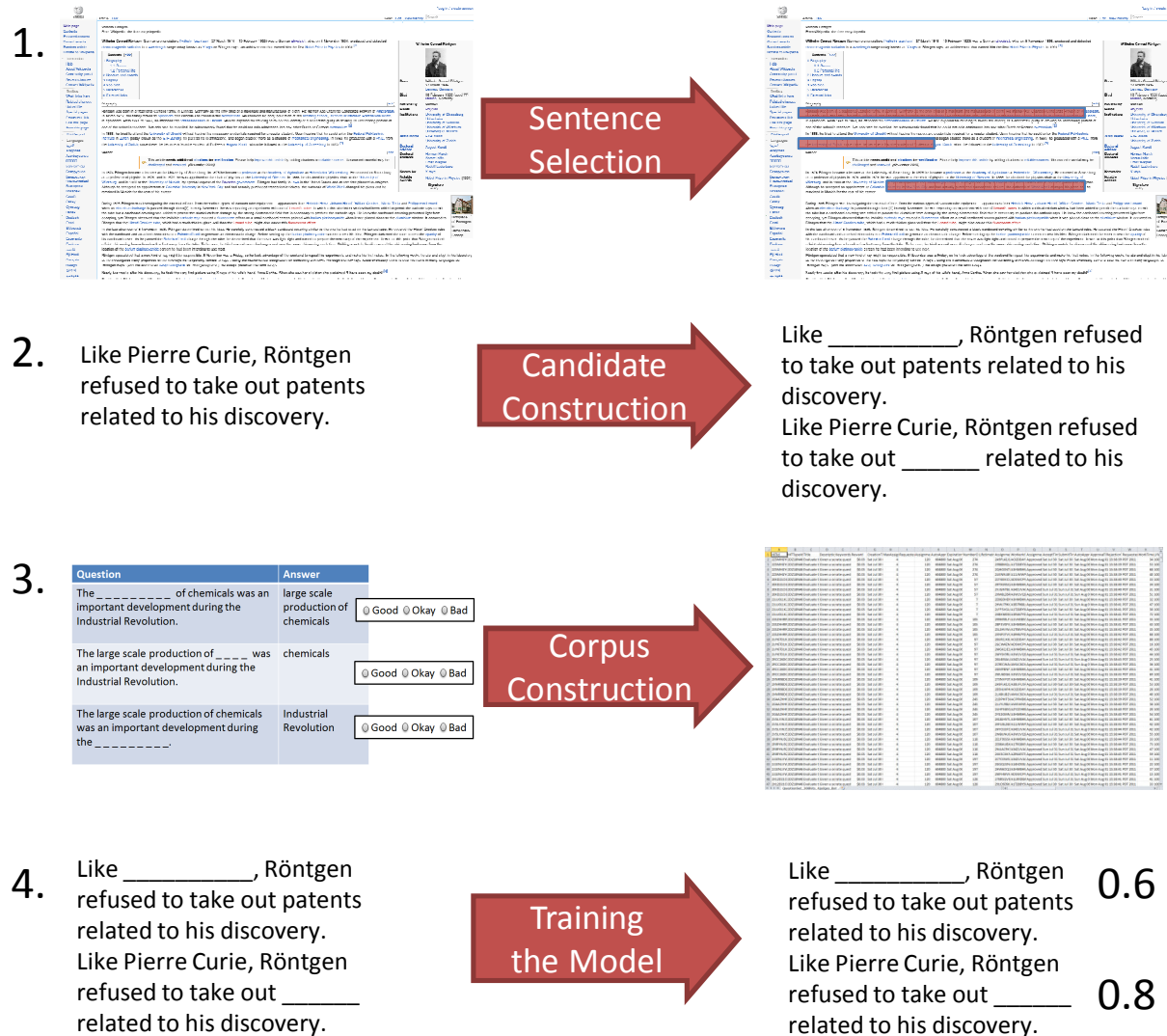
First Step of the Mastery Loop: Testing the Student

- Our goal: generate high quality questions from textbooks, web articles, or other source materials
 - First, we select the *most important parts* of the text to ask about
 - Then select the parts of those sentences that will make the *best questions*
 - Finally, create cloze (fill-in-the-blank)* questions from those parts
- The resulting questions can be useful to multiple audiences:
 - **Students**: for review and mastery
 - **Teachers**: as a “power tool” to help with creating exams

Question Generation: Related Work

- Wh-Questions
 - Autoquest (Wolfe, 1976)
 - Transformation rules (Mitkov and Ha, 2003)
 - Template-based generation (Chen et al., 2009)
 - Overgenerate-and-rank (Heilman and Smith, 2010)
 - QG-STEAC (Rus et al., 2010)
- Fill-in-the-blank (aka gap-fill & cloze) questions
 - Content-focused (Agarwal and Mannem, 2011)
 - Vocabulary and language learning (Pino et al., 2008)

Question Generation Overview



Candidate Construction

- Task: Given a sentence, generate a question that best covers the material in that sentence.
- Approach: Over-generate and rate candidates
 - Obtain constituency parse and SRL for each sentence
 - Create gap for each SR argument and each nested NP and AJP
 - Human judges to rate each candidate question

Candidate Construction Example

- Before Genghis Khan died, he assigned Ögedei Khan as his successor and split his empire into khanates among his sons and grandsons.

Candidate Construction Example

- Before Genghis Khan died, he assigned Ögedei Khan as his successor and split his empire into khanates among his sons and grandsons.
 1. Before _____ died, he assigned Ögedei Khan as his successor and split his empire into khanates among his sons and grandsons.
 2. Before Genghis Khan died, __ assigned Ögedei Khan as his successor and split his empire into khanates among his sons and grandsons.
 3. Before Genghis Khan died, he _____ Ögedei Khan as his successor and split his empire into khanates among his sons and grandsons.
 4. Before Genghis Khan died, he assigned _____ as his successor and split his empire into khanates among his sons and grandsons.
 5. Before Ghengis Khan died, he assigned Ögedei Khan as _____ and split his empire into khanates among his sons and grandsons.

Semantic Role Labels:

Pred A0 A1 A2 AM-TMP

Corpus Construction: HITs

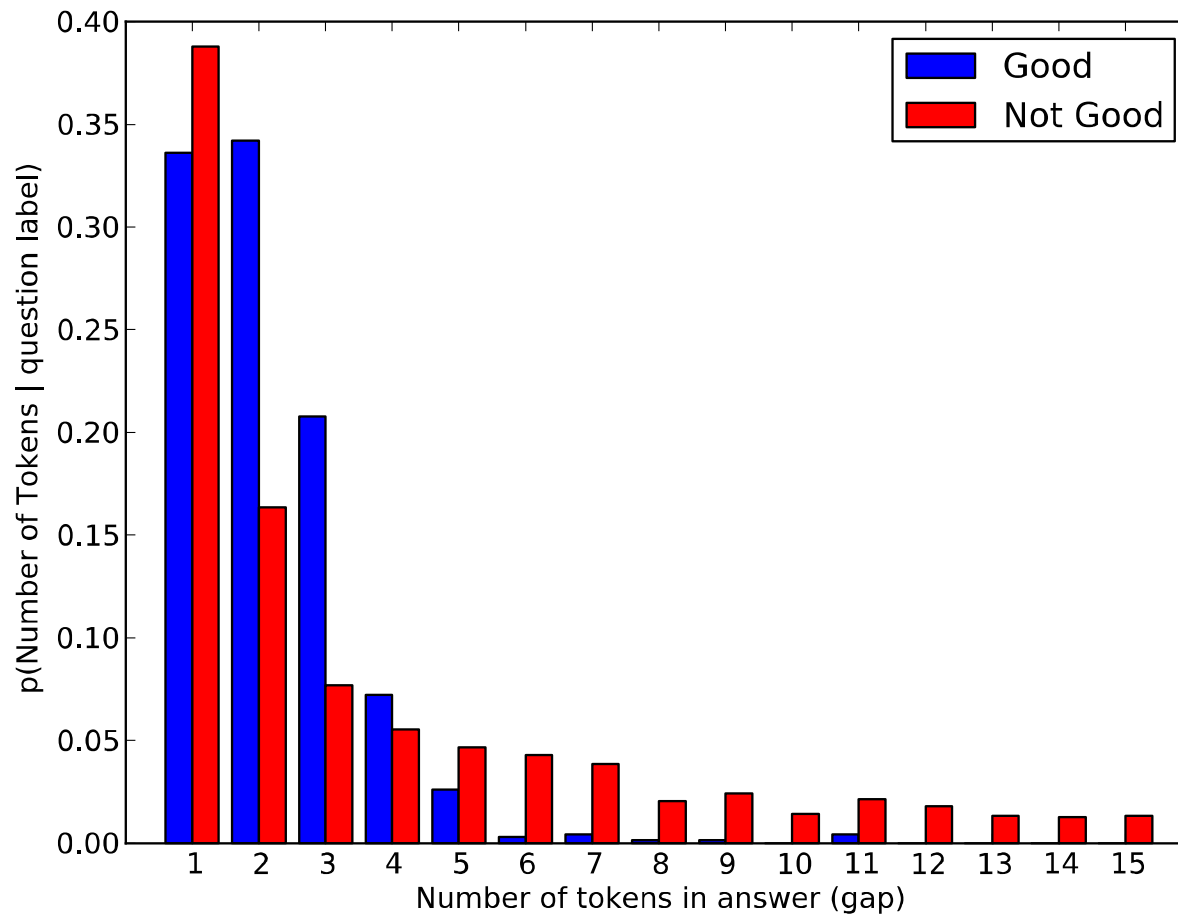
Question	Answer
_____ was an important development during the Industrial Revolution.	The large scale production of chemicals <input type="radio"/> Good <input checked="" type="radio"/> Okay <input type="radio"/> Bad
The large scale _____ of chemicals was an important development during the Industrial Revolution.	production <input type="radio"/> Good <input type="radio"/> Okay <input checked="" type="radio"/> Bad
The large scale production of chemicals was an important development during _____.	the Industrial Revolution <input checked="" type="radio"/> Good <input type="radio"/> Okay <input type="radio"/> Bad

Corpus Details

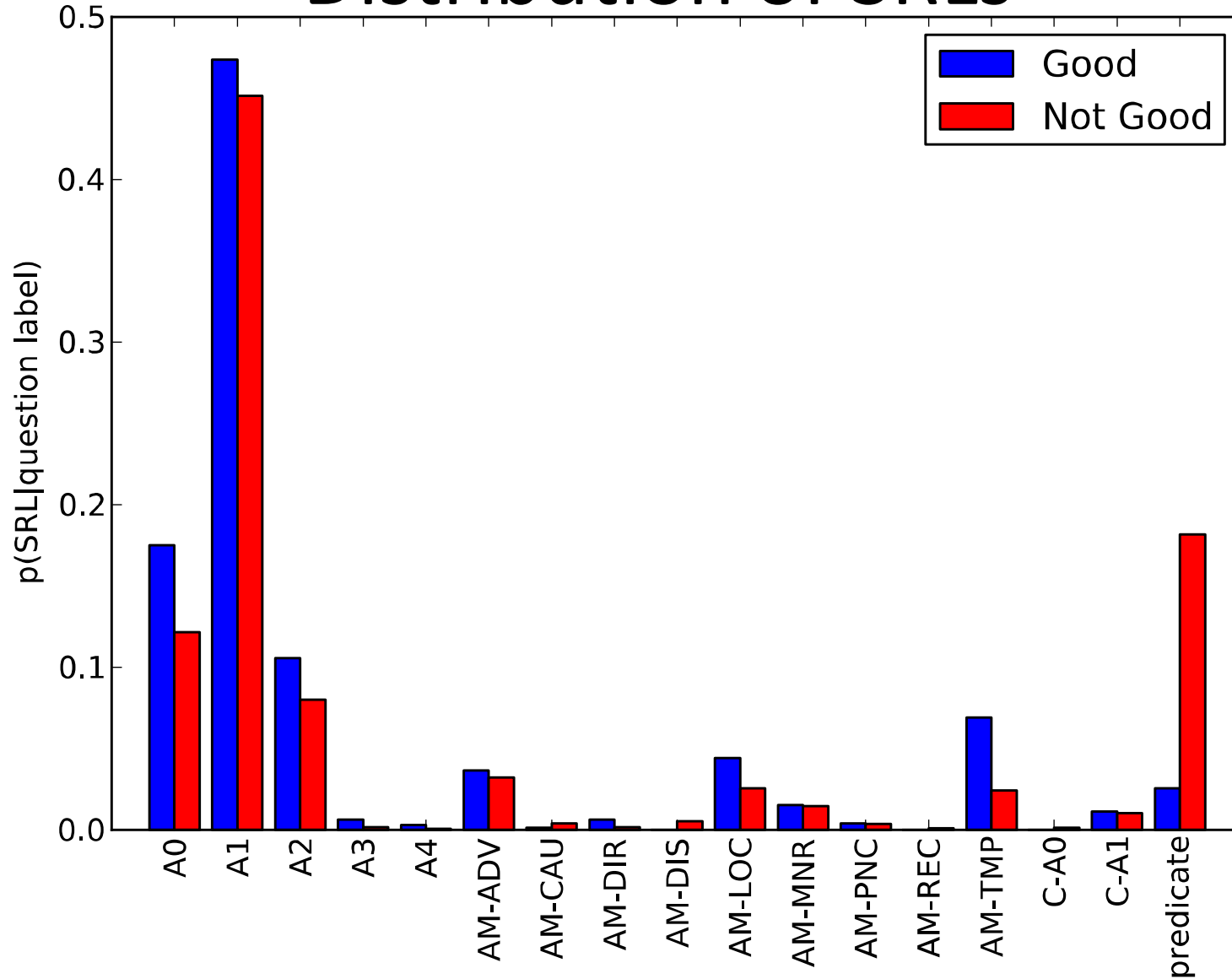
- 105 vital/popular Wikipedia articles
- Sentences: for each article, 10 from SumBasic + 10 from random sampling
- HITs: ~10 Questions / HIT, 4 judges/HIT
- 2252 Candidate Questions in total
- 85 unique judges
- Filtered workers and questions:
 - Eliminated 431 questions, Retained 1821 questions with highest agreement.
 - Of filtered questions 700 (38%) labeled *Good*

Corpus available at <http://research.microsoft.com/~sumitb/questiongeneration>

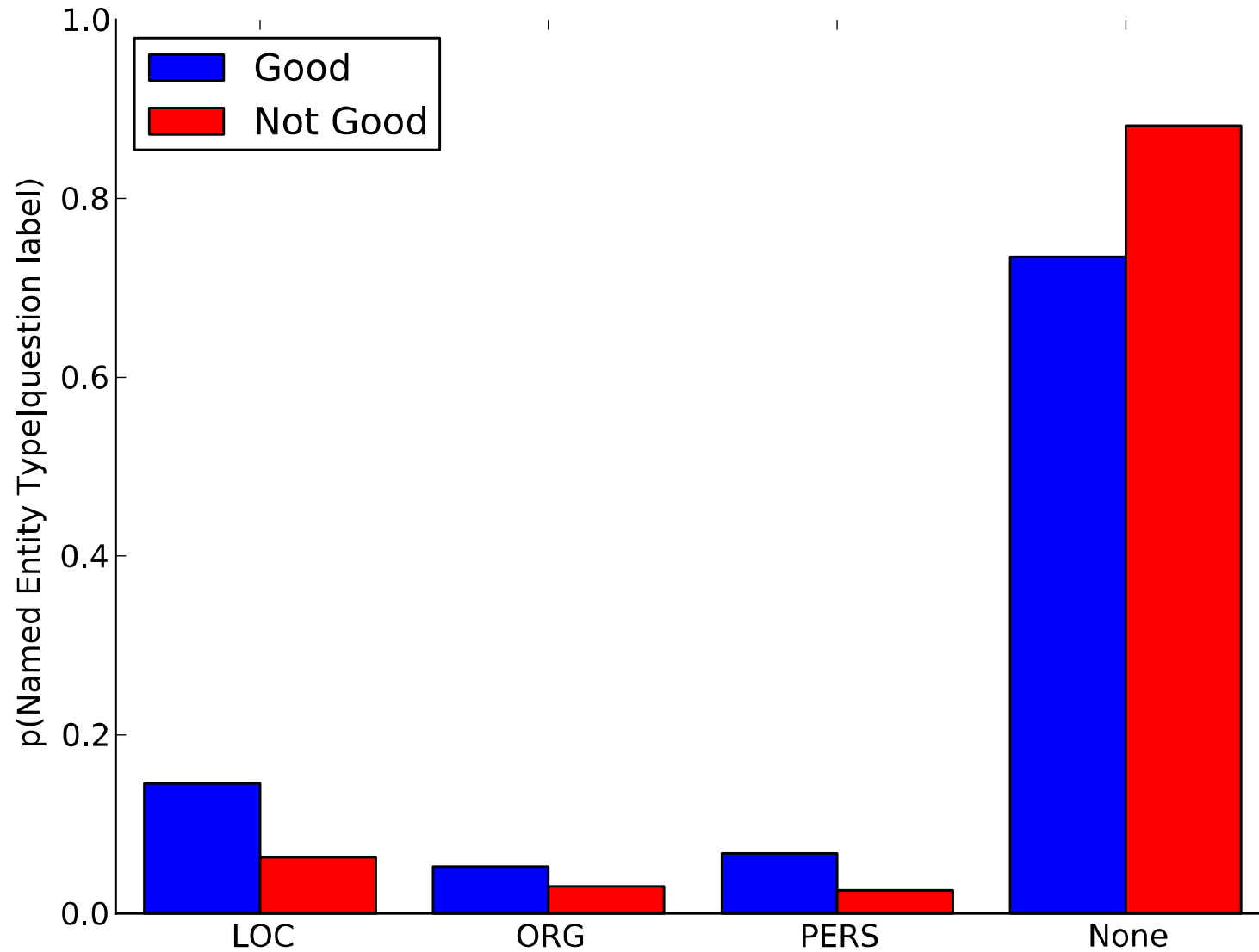
Examining the Corpus: Distribution of questions by gap length



Examining the Corpus: Distribution of SRLs



Examining the Corpus:



Training the Model:

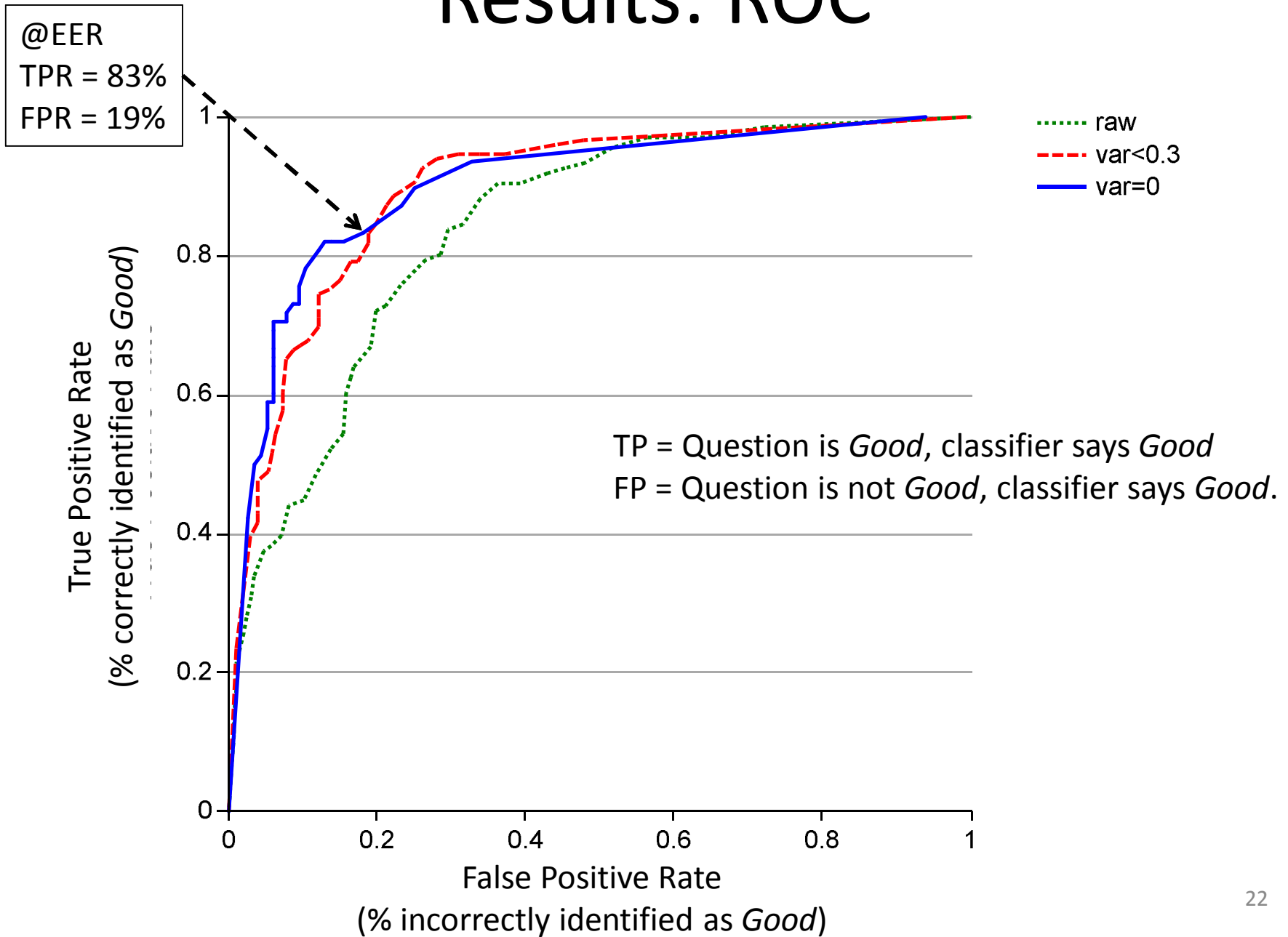
Gap selection as supervised learning

- Approach: Overgenerate and score:
 - Identify candidate blanks
 - Extract features from the sentence and the gap
 - Train/Evaluate ‘Good’ vs ‘not-Good’ question classifier.
 - For scoring use calibrated learner*
 - Logistic Regression + L2 Regularizer
 - Evaluation: 10-fold cross validation

Features

Feature Category	Number	Examples
Token Count	5	Num. tokens in sentence Num. overlapping tokens sentence:gap
Lexical	11	Gap [pronoun stopword abbrev. capital] density
Syntactic	112	POS tag before/after gap Gap bag of POS tags Gap syntactic parse depth Gap location relative to head verb (before/after)
Semantic	40	SRLs contained in gap SRL covering gap
Named Entity	11	Gap named entity density Gap named entity type frequency (LOC, ORG, PER) Sentence named entity frequency
Wikipedia Link	3	Gap link density Sentence link density
Total	182	

Results: ROC



Demo: Wikipedia Article on "Entropy": Original Text



The image shows a screenshot of a web browser displaying the Wikipedia article for "Entropy". The browser's address bar shows the URL "http://sumiz8/sho/qgder". The page title is "Wikipedia Article for 'entropy'". The article content includes the Wikipedia logo, navigation tabs (Article, Talk, Read, Edit, View history), and a search box. The main text defines entropy as a thermodynamic property and provides a detailed explanation of its role in thermodynamics, including a reference to the second law and the concept of disorder. An image of a glass of iced tea with a lemon slice is included, with a caption explaining it as an example of increasing entropy. A sidebar on the left contains various navigation links, and a "Generate Quiz" button is visible at the bottom.

Wikipedia Article for "entropy"

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export

Languages
العربية
বাংলা
Беларуская
Български
Boarisch
Bosanski
Brezhoneg
Català
Česky
Dansk
Deutsch
Eesti
Ελληνικά
Español
Esperanto

Create account Log in

Article Talk Read Edit View history Search

Entropy

From Wikipedia, the free encyclopedia

This article is about entropy in thermodynamics. For other uses, see [Entropy \(disambiguation\)](#).

For a generally accessible and less technical introduction to the topic, see [Introduction to entropy](#).

Entropy is a [thermodynamic property](#) that is the measure of a system's thermal energy per unit temperature that is unavailable for doing useful [work](#). Perhaps the most familiar manifestation of entropy is that, following the laws of thermodynamics, entropy of a closed [system](#) always increases and in [heat transfer](#) situations, heat energy is transferred from higher [temperature](#) components to lower temperature components. In thermally isolated systems, entropy runs in one direction only (it is not a reversible process). One can measure the entropy of a system to determine the energy not available for [work](#) in a [thermodynamic process](#), such as energy conversion, engines, or machines. Such processes and devices can only be driven by convertible energy, and have a theoretical maximum efficiency when converting energy to work. During this work, entropy accumulates in the system, which then [dissipates](#) in the form of waste [heat](#).

In classical thermodynamics, the concept of entropy is defined [phenomenologically](#) by the [second law of thermodynamics](#), which states that the entropy of an [isolated system](#) always increases or remains constant. Thus, entropy is also a measure of the tendency of a process, such as a chemical reaction, to be [entropically favored](#), or to proceed in a particular direction. It determines that [thermal energy](#) always flows spontaneously from regions of higher temperature to regions of lower temperature, in the form of [heat](#). These processes reduce the state of order of the initial systems, and therefore entropy is an expression of disorder or randomness. This is the basis of the modern microscopic interpretation of entropy in [statistical mechanics](#), where entropy is defined as the amount of additional information needed to specify the exact



Ice melting in a warm room is a common example of increasing entropy.^{[[note 1](#)]} described in 1862 by [Rudolf Clausius](#) as an increase in the [disgregation](#) of the water molecules in ice: order leading to disorder^[1]

Entropy articles
Introduction
History
Classical
Statistical

Thermodynamics

Generate Quiz

Demo: Wikipedia Article on "Entropy": Generated Questions



Quiz for article "entropy"

0.87 Henceforth, the essential problem in statistical thermodynamics , i.e. according to , has been to determine the distribution of a given amount of energy E over N identical systems.

0.87 One can measure the entropy of a system to determine the energy not available for in a thermodynamic process , such as energy conversion, engines, or machines.

0.86 It follows that heat will not flow from a colder body to a hotter body without the application of work (the imposition of) to the colder body.

0.84 This is the second law of .

0.83 linked entropy with a mathematical definition of irreversibility, in terms of trajectories and integrability.

0.82 In statistical mechanics, entropy is a measure of the number of ways in which a system may be arranged, often taken to be a measure of "" (the higher the entropy, the higher the disorder).

0.82 The second law is then a consequence of this definition and the fundamental postulate of .

0.81 Importantly, it makes no reference to the microscopic nature of .

0.78 This condition is known as the " heat death of " .

0.75 The concept of entropy arose from Rudolf Clausius 's study of .

0.71 So more heat is given off to the cold reservoir than in .

0.68 The information entropy H for is

0.63 Using this concept, in with the density matrix he extended the classical concept of entropy into the quantum domain.

0.61 where p is the probability of a system's being in a particular microstate, given that it is in a particular macrostate, and is constant.

0.59 So in this example, the entropy of the system increases, whereas of the surroundings decreases.

0.58 From a macroscopic perspective, in classical thermodynamics the entropy is interpreted as a state function of a thermodynamic system : that is, a property depending only on the current state of , independent of how that state came to be achieved.

0.56 If we denote the entropies by for the two states, then the above inequality can be written as a decrease in .

0.54 The entropy of has decreased as some of its energy has been dispersed to the ice and water.

0.42 In thermally isolated systems, entropy runs in only (it is not a reversible process).

Show Answers

Demo: Wikipedia Article on "Entropy": Answers



Quiz Answers for Article "entropy"

Question: Henceforth, the essential problem in statistical thermodynamics , i.e. according to _____, has been to determine the distribution of a given amount of energy E over N identical systems.
Correct Answer: [Erwin Schrödinger](#)
Your Answer: [Newton](#)

Question: One can measure the entropy of a system to determine the energy not available for ___ in a thermodynamic process , such as energy conversion, engines, or machines.
Correct Answer: [work](#)
Your Answer: [work](#)

Question: It follows that heat will not flow from a colder body to a hotter body without the application of work (the imposition of ___) to the colder body.
Correct Answer: [order](#)
Your Answer: [order](#)

Question: This is the second law of _____ .
Correct Answer: [thermodynamics](#)
Your Answer: [thermodynamics](#)

Question: _____ linked entropy with a mathematical definition of irreversibility, in terms of trajectories and integrability.
Correct Answer: [Carathéodory](#)
Your Answer: [Newton](#)

Question: In statistical mechanics, entropy is a measure of the number of ways in which a system may be arranged, often taken to be a measure of " ___ " (the higher the entropy, the higher the disorder).
Correct Answer: [disorder](#)
Your Answer: [disorder](#)

Question: The second law is then a consequence of this definition and the fundamental postulate of _____ .
Correct Answer: [statistical mechanics](#)
Your Answer: [conservation of energy](#)

Question: Importantly, it makes no reference to the microscopic nature of ____ .
Correct Answer: [matter](#)
Your Answer: [matter](#)

Question: This condition is known as the " heat death of _____ " .
Correct Answer: [the Universe](#)
Your Answer: [the universe](#)

Question: The concept of entropy arose from Rudolf Clausius 's study of _____ .
Correct Answer: [the Carnot cycle](#)
Your Answer: [heat](#)

Question: So more heat is given off to the cold reservoir than in _____ .
Correct Answer: [the Carnot cycle](#)
Your Answer: [the warm one](#)

Question: The information entropy H for _____ is
Correct Answer: [equal probabilities](#)
Your Answer: [a system](#)

Example Results: False Positives

Raters considered these bad

Question	Answer	SystemScore
In 1821 , the Greeks declared __ on the Sultan.	war	0.732
This includes greeting others with " as-salamu `alaykum " ("peace be unto __"), saying bismillah ("in the name of God ") before meals, and using only the right hand for eating and drinking.	you	0.907
Not only is there much ice atop _____, the volcano is also slowly being weakened by hydrothermal activity.	the volcano	0.790

Example Results: False Negatives

Raters considered these good

Question	Answer	System Score
Caesar then pursued Pompey to Egypt, where Pompey was soon _____ .	murdered	0.471
About 7.5% of world sea trade is carried via the canal _____ .	today	0.119
Asante and Dahomey concentrated on the development of "legitimate commerce" in _____ , forming the bedrock of West Africa's modern export trade.	the form of palm oil , cocoa , timber and gold	0.029

Question Generation on Plain Text

Question Generation Demo

This is a demonstration of the question generation technology developed at MSR and detailed in the paper [Mind the Gap: Learning to Choose Gaps for Question Generation](#). It also serves as an example of how to use the [MSR question generation web service](#) as a RESTful service from Javascript using JQuery (a SOAP endpoint is also available). To try it out, enter or paste text in the box below and then hit the "Generate Questions" button. Please contact Sumit Basu (sumitb) and Lucy Vanderwende (lucyv) with additional questions.

Source Text

We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pursuit of Happiness. That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed. That whenever any Form of Government becomes destructive of these ends, it is the Right of the People to alter or to abolish it, and to institute new Government, laying its foundation on such principles and organizing its powers in such form, as to them shall seem most likely to effect their Safety and Happiness. Prudence, indeed, will dictate that Governments long established should not be changed for light and transient causes; and accordingly all experience hath shewn that mankind are more disposed to suffer, while evils are sufferable than to right themselves by abolishing the forms to which they are accustomed. But when a long train of abuses and usurpations, pursuing invariably the same Object evinces a design to reduce them under absolute Despotism, it is their right, it is their duty, to throw off such Government, and to provide new Guards for their future security. Such has been the patient sufferance of these Colonies; and such is now the necessity which constrains them to alter their former Systems of Government. The history of the present King of Great Britain is a history of repeated injuries and usurpations, all having in direct object the establishment of an absolute Tyranny over these States. To prove this, let Facts be submitted to a candid world.

Generate Questions

Then What?

- For the student case:
 - Because they are using this system to help them study, *they can grade their own answers.*
 - Adaptation: we can then *adapt the reading material based on their performance*, to focus on those areas where they need the most work
- Expanding the types of possible questions
 - Generating high-level concept questions covering larger spans of text
 - Well-formed Wh- questions from identified spans

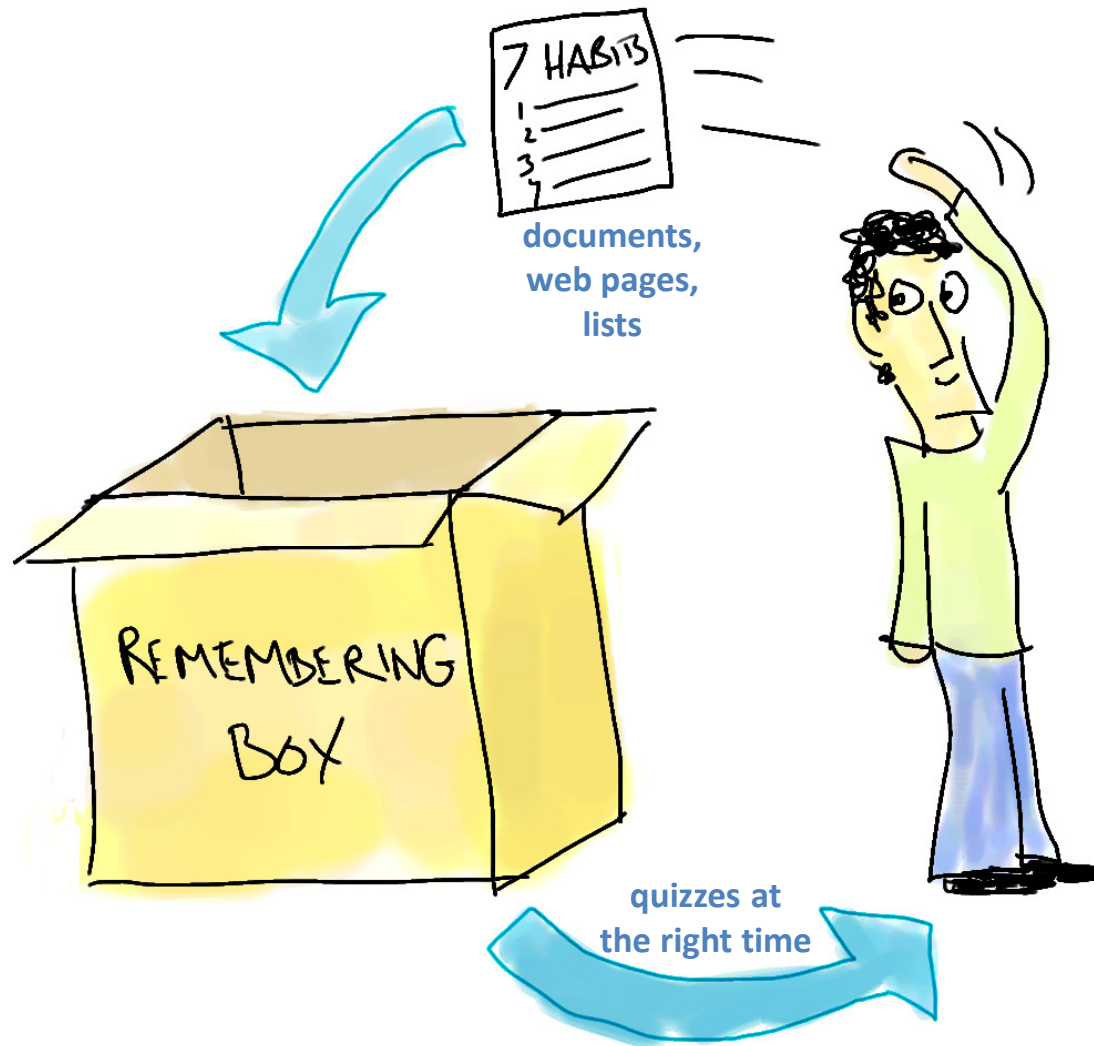
Grading Questions

- How can we grade fill-in-the-blank questions?
- Can we do it quickly, cheaply, accurately?
- Gave 1280 sections to Turkers (320x4 judges), 5 q's each (6400 total)
 - 1: turkers read section
 - 2: we hid the section and gave them the quiz
 - 3: they saw the true answer and their own, asked to self-grade
- 984 items graded by two experts (Sumit/Lucy)
- 911 items where experts gave the same grade
- We also distributed first 1000 questions to other Turkers to grade
- Next step – a calibrated automatic means of grading that can shunt to Turkers

Table 1: Agreement of various methods with experts on the 911 question/answer pairs where both experts agreed on the grade

Method	Agreement	More Harsh	More Lenient
Self Grading	93.5%	4.5%	2.0%
Turker Grading	95.4%	2.4%	2.2%
String Match	79.1%	20.9%	0.0%

Lifelong Memorization



Goal: help you master and refresh important content for a lifetime

Axis 2: Improving Coverage

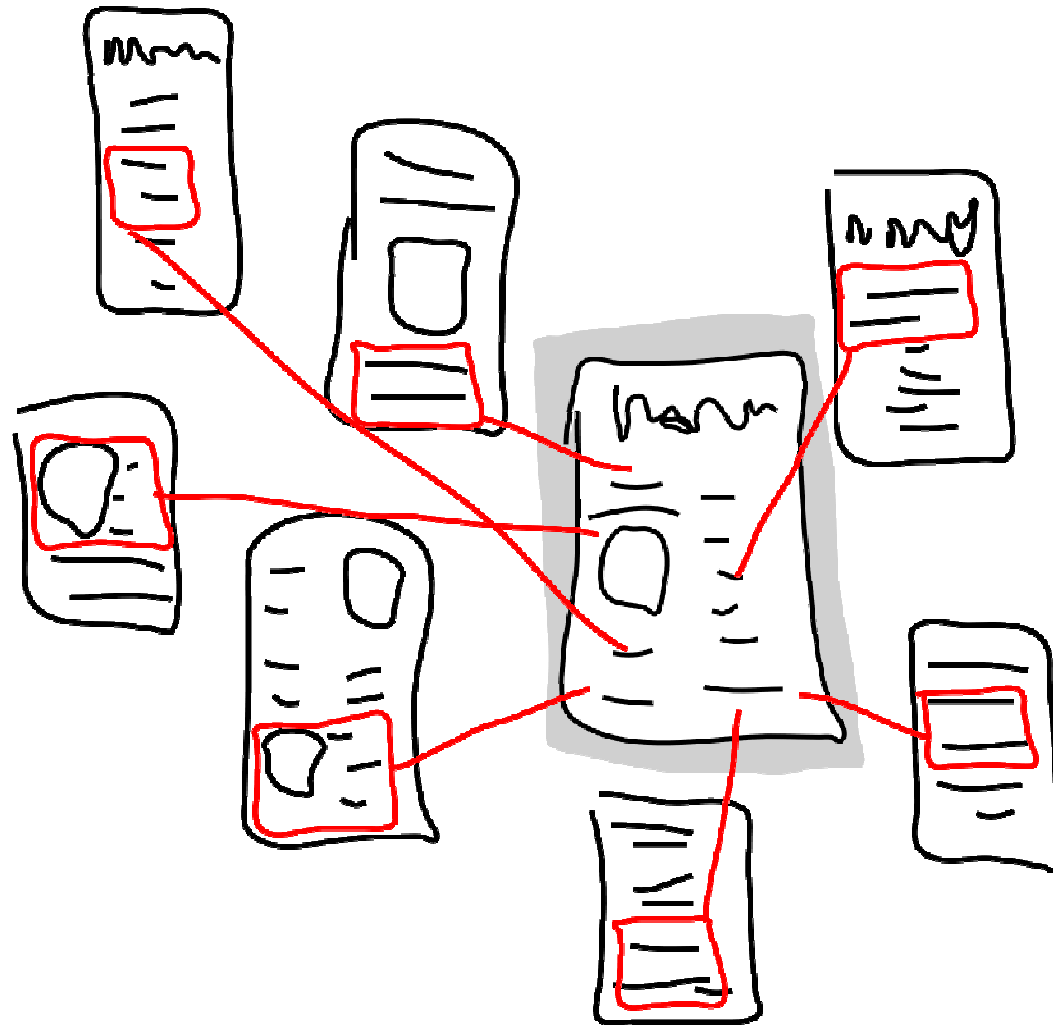
- When we're reading to learn, **how do we know when we've read enough?**
- **How do the set of all relevant documents connect** to what we've chosen to read?
- How do we connect what we're reading now to **what we've read in the past?**
- In order to learn more, **what should we read next** after reading this document?

How Do You Learn from a Document Collection?

“bing, I need to learn more about anemia...”



Finding Connections



Axis 3: Improving Engagement

- How can we help people use their reading time more effectively?
- How can we get people to read more?
- Can we make long reading tasks less daunting?
- Can we help readers reflect on their reading progress in a topic area?

This space reserved for
Audience-Generated Questions