Pearson Knowledge Technologies, PKT, is a subsidiary of Pearson, Plc, London.
Underlying technology developed at Bell Labs, Bell Communications Research the University of Colorado, and New Mexico State University.
Currently ~20 professionals [cognitive psychologists, computer scientists, statisticians software engineers, misc.]
Tom Landauer
EVP Research, Pearson Knowledge Technologies
Prof. Emeritus and Senior Research Associate,
University of Colorado, USA

First European Workshop
LSA for Education
The Open University, Heerlen, The Netherlands, March 29-30, 2007
Simon Dennis
Scott Dooley
Susan Dumais
Peter Foltz
George Furnas
Walter Kintsch
Darrell Laham
Michael Littman
Karen Lochbaum
Dian Martin
Rob Oberbreckling
Bill Oliver
Bob Rehder
Mark Rosenstein
Lynn Streeter
Jose Quesada

4/20/2007
Outline

1. What’s so good about LSA?

2. A sample of ways in which PKT has applied it to education
   - Automatically assessing essays content (fairly complete description)
   - Tutoring reading and substantive writing practice
   - Facilitating knowledge acquisition of from text
   - Aligning text and tests with content
   - Enhancing collaborative learning
   - Matching people with instruction and occupations
How it works

\[ \cos \Theta = \frac{\sum_{i=1}^{m} a_i b_i}{\sqrt{\sum_{i=1}^{m} a_i^2} \sqrt{\sum_{i=1}^{m} b_i^2}} \]
An Intuitive Explanation of Latent Semantic Analysis

- **Humans** learn word meanings and how to combine them into passage meaning through experience with ~paragraph unitized verbal environments.
- They don’t remember all the separate words of a passage; they remember its overall gist or meaning.

- **LSA** learns by “reading” ~paragraph unitized texts that represent the environment.
- It doesn’t remember all the separate words of a text it; it remembers its overall gist or meaning.
The most wonderful thing it can do

_______ is very probably Holland’s most beautiful city

The loveliest town in the Netherlands may well be_______

Similarity (\(\sqrt{\text{Cos}}\)) = .82
How did that happen?

Meaning of paragraph ≈
 meaning of $w_1$ + meaning of $w_2$ + … + meaning of $w_n$

Made true for a huge representative corpus of meaningful texts, justifying its likely validity everywhere
What else does it need?

Some approximate rules of thumb:

A corpus of authentic, attested, relevant text with
> 20,000 word types that are each used in ~10 different
~paragraphs and > ~20,000 paragraphs

About small ones…why??
Test: ask, can you learn a language, say Dutch, well
enough for the task based on studying corpus alone?

If computational…

Reduce to 200-500 dimensions, no less, no more,
presumably the natural dimensionality
Imagine I give you 500 paragraphs of Swahili to study…
What does it not usually want?

Stemming
Stop lists
Frequency cutoffs
Low or high D
Interpreting dimensions
The important thing that LSA tells is not how often words have or will appear in the same contexts…

…it is that it tells the degree to which words will combine to produce the same effects on passage meanings.
So, stamp out “co-occurrence” !!

If you want to know about Lemaire and Denhier’s experiment, ask me later
Over 99% of word-pairs whose similarity is induced never appear together in a paragraph.
Correlations (r) with LSA cosines over 10,000 random wd-wd pairs:

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time two words co-occur in same paragraph (log both)</td>
<td>0.35</td>
</tr>
<tr>
<td>Time two words occur in separate paragraphs (log A only + log B only)</td>
<td>0.30</td>
</tr>
<tr>
<td>Contingency measures:</td>
<td></td>
</tr>
<tr>
<td>Mutual information</td>
<td>0.05</td>
</tr>
<tr>
<td>Chi-square</td>
<td>0.10</td>
</tr>
<tr>
<td>Joint/expected p(A&amp;B) / (p(A)*p(B))</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Variance not accounted for by frequency or contingency of joint occurrence: About 85% or more.
Beats ordinary vector space

• **Words:**

<table>
<thead>
<tr>
<th>Word overlap</th>
<th>LSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctor—Doctor</td>
<td>1.0</td>
</tr>
<tr>
<td>Doctor—Physician</td>
<td>0.0</td>
</tr>
<tr>
<td>Doctor—Surgeon</td>
<td>0.0</td>
</tr>
<tr>
<td>Surgeon--Physician</td>
<td>0.0</td>
</tr>
</tbody>
</table>

• **Passages:**

1. Doctors operate on patients.
2. Physicians do surgery.

Keywords 0, LSA 90%
Average of all 12 word-to-word = 65%
How it does without syntax:
Potential information in a paragraph.

100,000 word vocabulary
average sentence = 20 words
from word combinations,
\[ \log_2(105*20) = 1,660 \text{ bits} \]
from word order,
\[ \log_2 (20!) = 305 \text{ bits} \]

\[ \therefore 84\% \text{ in word choice} \]
Educated adult understands ~100,000 word forms
an average sentence contains 20 tokens.
thus 100,000^{20} possible combinations of words in a sentence
∴ maximum of log_2 100,000^{20} = 332 bits in word choice alone.
20! = 2.4 \times 10^{18} possible orders of 20 words
∴ maximum of 61 bits from order of the words.
332/(61 + 332) = 84\%. word choice
In scoring two large sets of GMAT with LSA alone (none of the other variables usually used)
mutual information in scores of LSA and one expert human grader averaged .81 mutual information in scores of two human graders .90.

∴ only .09/.90 = 10% of the information that the humans extracted was missed by LSA
Outline

1. What’s so good about LSA?

2. Examples of ways to apply LSA to education
   - Automatic essay scoring (a fairly complete description)
   - Reading and substantive writing practice
   - Facilitating knowledge acquisition of from text
   - Aligning text and tests with content
   - Enhancing collaborative learning
   - Matching people with instruction and occupations
   - Doing these in other languages than English
The Intelligent Essay Assessor (IEA)

Pre-trained on >> 50,000 paragraphs
of representative text
Near neighbor measure

human score: 3
human score 4
human score 3
human score 5
human score 4
human score 5
Score ????
human score 5
human score 5
Example reliability
Examples of new IEA components:

Text-specific Bayesian 5-grams for backward and forward word order

Contextual choice and distribution of words

Semantic coherence

Off topic, oddity and plagiarism detection

….maybe entropy, SVM…

<= ~10 best and intuitively defensible variables
How might a college education help you? Describe in detail three ways in which a college education can contribute to long-term life success.

Please answer this question in 250-350 words.

A person who has acquired a college level degree has many advantages over a person who does not have a degree. A college education is helpful because it can make a person more competitive in the job market. A person with a college degree has high employability and high earning potential, developed a variety of skills that can be used in the workplace, and also has acquired lifelong learning skills that can be used forever.

When a person graduates from college, statistics show that they can compete for a higher income than people.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education gives you tools for lifelong learning.</strong> You learn facts while you are in school, but more importantly, you learn how to think. While some of the facts and figures you learn today may not apply to the world of tomorrow, your ability to think will be useful always, in everything you do.</td>
<td>This is a topic you didn't address.</td>
</tr>
<tr>
<td><strong>Education improves your quality of life.</strong> Income and employment get a boost from education. The Digest of Education Statistics 1996 reports that income levels rise as educational levels rise. Figure 1-1 shows average income levels for different levels of educational attainment. Figure 1-2, also from a report in the Digest, shows how unemployment rates decrease as educational levels rise.</td>
<td>👍 You covered this topic well.</td>
</tr>
<tr>
<td><strong>Education expands your self-concept.</strong> As you rise to the challenges of education, you will discover that your capacity for knowledge and personal growth is greater than you imagined. As your abilities grow, so do opportunities to learn and do more in class, on the job, and in your community.</td>
<td>👍 You addressed this topic, though you might have said more.</td>
</tr>
<tr>
<td><strong>Education enlarges your possibilities.</strong> Education gives you a base of choices and increased power, as shown in Figure 1-3. First, through different courses of study, it introduces you to more choices of career and life goals. Second, through the training you receive, it gives you more power to achieve the goals you choose. For example, while taking a writing class, you may learn about careers in journalism. This experience may lead you to take a class in journalistic writing that teaches you about reporting. Down the road, you may decide to work on a newspaper and to make journalism your career. Looking back, you realize that two classes you took in college changed the course of your life.</td>
<td>This is a topic you didn't address.</td>
</tr>
<tr>
<td><strong>Education improves your employability and earning potential.</strong> Learning additional skills raises your competency so you can fulfill the requirements of higher-level jobs. In addition, having a college degree makes an impression on potential employers and makes you eligible for higher-salaried positions.</td>
<td>👍 You covered this topic well.</td>
</tr>
<tr>
<td><strong>Education makes you a well-rounded person.</strong> As it widens your understanding about what is possible in the world, education increases your awareness and appreciation of areas that affect and enrich human lives, such as music, art, literature, science, politics, and economics.</td>
<td>👍 You addressed this topic, though you might have said more.</td>
</tr>
<tr>
<td><strong>Education affects both community involvement and personal health.</strong> Education helps to prepare individuals for community activism by helping them understand political, economic, and social conditions. Education also increases knowledge about health behaviors and preventive care. The more education you have, the more likely you are to practice healthy habits in your daily life and to make informed decisions.</td>
<td>👍 You addressed this topic, though you might have said more.</td>
</tr>
</tbody>
</table>
An example of plagiarism detection

Mainframes are primarily referred to large computers with rapid, advanced processing capabilities that can execute and perform tasks equivalent to many Personal Computers (PCs) machines networked together. It is characterized with high quantity Random Access Memory (RAM), very large secondary storage devices, and high-speed processors to cater for the needs of the computers under its service.

Mainframes usually are referred to computers with fast, advanced processing capabilities that could perform by itself tasks that may require a lot of Personal Computers (PC) Machines. Usually mainframes would have lots of RAMs, very large secondary storage devices, and very fast processors to cater for the needs of those computers under its service.
Recent IEA reliabilities and validities

**MetaMetrics experiment:**
2,579 essays 12 prompts, grades 4,6,8,10,12
IEA-human
  .76-.90  mean = .85

Average effect size of two school grades
  Human scores: .42
  IEA scores: .42
Recent IEA reliabilities and validities

Prentice Hall commercial scoring program

81 prompts, grades 6-12, 33,205 essays

<table>
<thead>
<tr>
<th></th>
<th>Human-Human</th>
<th>IEA-human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.86</td>
<td>0.89</td>
</tr>
<tr>
<td>Exact agreement</td>
<td>62.1%</td>
<td>61.3%</td>
</tr>
<tr>
<td>+ adjacent</td>
<td>97.7%</td>
<td>98.1%</td>
</tr>
</tbody>
</table>
Traits
Word, Focus, Organization, Voice, Mechanics…
  human and IEA ~.5-.6

IEA sans LSA, , no content
  ~ 11% less
Outline

1. What’s so good about LSA?

2. Examples of ways to apply LSA to education
   - Automatic essay scoring (a fairly complete description)
   - Reading and substantive writing practice
   - Facilitating knowledge acquisition of from text
   - Aligning text and tests with content
   - Enhancing collaborative learning
   - Matching people with instruction and occupations
   - Doing these in other languages than English
Summary Street

- Student reads a curriculum relevant text
- Copies or types a summary ~20% as long
- System evaluates how well each section or topic is covered
- Student revises until a system or teacher-determined criterion is met
Summary Street

• also:
  – Shows redundant or irrelevant sentences
  – Checks spelling with optional help
  – Gives limited grammar feedback
  – Flags excessive copying from source
Write Your Summary

Your last summary for Wind as Energy is below.

User: student1  Text: Wind as Energy

Your summary should be 100 to 250 words long. Type your summary in the box below.

Wind is caused by the uneven heating of the earth’s surface. Since the earth’s surface is made up of land, desert, water, and forest areas, the surface absorbs the sun’s radiation differently. During the day, air above the lands heat more quickly than air over water.

More than 5,000 years ago ancient Egyptians used the wind power to sail ships on the Nile river.

Today, wind machines will still use blades to get the wind's kinetic energy. Horizontal-axis machines use a tail to keep the rotor facing the wind but vertical axis turbines catch the wind better.

Wind energy is an important alternative energy source for the future. The cost of producing wind energy has dropped dramatically in the last two decades. In 1978, the wind cost 30 cents per kWh, but now costs less than 5 cents. In comparison, new coal plants produce electricity for about 8 cents per kWh.

The student can check spelling or get immediate content feedback.
The bars show how well the summary covered the sections. The summary length bar shows whether the summary is an appropriate length.

Feedback is also provided on spelling, irrelevance, redundancy, and plagiarism.
Student One, your summary is shown below. Sentences that may be irrelevant are marked in red.

If you agree that the sentences that are marked are irrelevant, either delete them or combine their information with other sentences in your summary.

Note: Spelling errors can cause good sentences to be marked as irrelevant, so be sure to check your spelling.

Your summary:

Wind is caused by the uneven heating of the earth's surface by the sun. Since the earth's surface is made up of land, desert, water, and forest areas, the surface absorbs the sun's radiation differently. During the day, arid lands heat more quickly than air over water.

More than 5,000 years ago ancient Egyptians used the wind power of the Nile river.

Today's wind machines will still use blades to get the winds' kinetic energy. Horizontal-axis machines use a tail to keep the rotor facing the wind but vertical axis catches wind better.

Wind energy is an important alternative energy source for the future. The cost of producing electricity from the wind has dropped dramatically in the last two decades. Electricity generated by the wind cost 30 cents per kWh in 1975, but now costs less than five cents per kWh. In comparison, new coal plants produce electricity at four cents per kWh.
This report shows detailed performance data on all texts summarized by a student in a class.
IEA + Tutorial

Better summaries as judged by teachers:
Middle School Students

- Content while using
- Quality while using
- Later on their own

- Write-to-learn
- Word processor

Time (min) students spent writing on their own
Difference in scores grade 4-6, 6-8, 8-10

By IEA > by mean of two human scores
Outline

1. What’s so good about LSA?

2. Examples of ways to apply LSA to education
   - Automatic essay scoring (a fairly complete description)
   - Reading and substantive writing practice
   - *Facilitating knowledge acquisition of from text
   - Aligning text and tests with content
   - Enhancing collaborative learning
   - Doing these in other languages than English
SuperManual and auto-autodidact

• Find and understand what, how and how-to information faster
• Combines LSA meaning-based with keyword search
• Maintains context constantly
• Text selection to paragraph search
320 ELECTRIC POWER DISTRIBUTION SYSTEMS

1. DESCRIPTION OF ELECTRICAL SYSTEMS AND EQUIPMENT

320-1.6 DISTRIBUTION SYSTEM EQUIPMENT

320-1.6.6 FUSES.

320-1.6.6 FUSES. A fuse consists of a metal conductor inserted into a tube of glass or other insulating material, that melts when the current through the conductor exceeds the rated level, thus opening the circuit. Metal ferrules at each end of the fuse make contact with fuse clips or contacts in the carriage of a screw-type fuse holder. Fuses are used as protective devices in power and lighting circuits. They are in some user equipment, where their performance is preferred over that of a circuit breaker or their use is considered more economical. Motors rated up to 7-1/2 horsepower (hp) are often supplied from group control centers having 30-ampere fuses. Fuses are not used in 450V circuits supplying motors in excess of 7-1/2 hp. Unlike circuit breakers, a fuse must be replaced when it fails. See NSTM Chapter 300 for directions on removing and replacing fuses.

320-1.6.6.1 Fuse Characteristics. There are three types of fuses with one of the following characteristics:

a. Characteristic A - normal blowing

b. Characteristic B - time lag (slow blow)

c. Characteristic C - very high (100,000 amperes) interrupting capacity

C fuses are used where the available fault current exceeds the 10,000 amperes maximum interrupting capacity of A or B fuses.

WARNING
320 ELECTRIC POWER DISTRIBUTION SYSTEMS
1. DESCRIPTION OF ELECTRICAL SYSTEMS AND EQUIPMENT
320-1.6 DISTRIBUTION SYSTEM EQUIPMENT
320-1.6.5 CIRCUIT BREAKERS.

320-1.6.5 CIRCUIT BREAKERS. Circuit breakers are mounted as an integral part of switchboards, switchgear groups, and distribution panels. Circuit breakers used on naval ships are classified ACB, AQB, AQB-LF, AQB-LL, NQB, ALB, and NLB. All of these types are open case, molded case, and insulated case circuit breakers.

320-1.6.5.1 ACB Type. ACB circuit breakers have an open, metallic frame construction that may be either manually or electrically operated. Electrical operation allows the circuit breaker to be used for remote control operation as well as for protection. ACB circuit breakers are commonly used to connect ship service and emergency generators to the power distribution system. They are also used on bus ties, some shore power connection circuits, and some larger feeder circuits.
Correlation of learning content to learning standards

- Content must correspond to appropriate State or School District Learning Objectives

- **Standards Seeker** measures semantic similarity of standards/learning objectives to educational material
  - Textbooks
    - Paragraphs, pages, sections
  - Lesson plans
  - Test items
  - Other standards/learning objectives
MAB.1.3.1.8.1
Strand B: Measurement
Standard 1: The student measures quantities in the real world and uses the measures to solve problems.

"Benchmark MAB.1.3.1: The student uses concrete and graphic models to derive formulas for finding perimeter, area, surface area, circumference, and volume of two- and three-dimensional shapes, including rectangular solids and cylinders."

8-1. uses concrete and graphic models to explore and derive formulas for surface area and volume of three-dimensional regular shapes, including pyramids, prisms, and cones"

MAB.1.3.1.8.2
Strand B: Measurement
Standard 1: The student measures quantities in the real world and uses the measures to solve problems.

"Benchmark MAB.1.3.1: The student uses concrete and graphic models to derive formulas for finding perimeter, area, surface area, circumference, and volume of two- and three-dimensional shapes, including rectangular solids and cylinders."

8-2. solves and explains real-world problems involving surface area and volume of three-dimensional shapes

MAB.1.3.2.8.1
Strand B: Measurement
Standard 1: The student measures quantities in the real world and uses the measures to solve problems.

"Benchmark MAB.1.3.2: The student uses concrete and graphic models to derive formulas for finding rates, distance, time, and angle measures."

8-1. applies formulas for finding rates, distance, time and angle measures"

MAB.1.3.2.8.2
Strand B: Measurement
Outline

1. What’s so good about LSA?

2. Some ways we have applied it to education
   - Automatic essay scoring (with a little detail)
   - Reading and substantive writing practice
   - Facilitating acquisition of knowledge from text
   - Aligning text and tests with content
   - Enhancing collaborative learning
   - Matching people with instruction and occupations
Outline

1. What’s so good about LSA?

2. Some ways we have applied it to education
   • Automatic essay scoring (with a little detail)
   • Reading and substantive writing practice
   • Facilitating acquisition of knowledge from text
   • Aligning text and tests with content
   • *Enhancing collaborative learning
   • Matching people with instruction and occupations
Knowledge Post

A collaborative teaching and learning environment

Developed for military training and distributed decision making

But we think promising for teacher professional development and ongoing improvement
Knowledge Post

Automatic

• Monitoring of activity
• Assessment of individual contributions, group progress
• Semantic search:
  – Prior postings
  – Repositories
• Mouse-over summaries
• Expert interjections
A user searches for comments related to the ‘students’ summaries are too short’ comment. The most similar comments are displayed.

**Subject**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Similarity (0-100)</th>
<th>Find Related</th>
<th>Author</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>in my experience...</td>
<td>(53)</td>
<td>Notes, References</td>
<td>teacherTom</td>
<td>09/14/04 05:22 PM</td>
</tr>
<tr>
<td>summarization activities to try</td>
<td>(45)</td>
<td>Notes, References</td>
<td>Angela</td>
<td>09/14/04 05:25 PM</td>
</tr>
<tr>
<td>Have you tried WriteToLearn?</td>
<td>(43)</td>
<td>Notes, References</td>
<td>Ms. Streeter</td>
<td>09/14/04 05:11 PM</td>
</tr>
<tr>
<td>Practice comprehension skills more</td>
<td>(40)</td>
<td>Notes, References</td>
<td>Mr. Jack</td>
<td>09/14/04 05:24 PM</td>
</tr>
<tr>
<td>Teaching Tips: Summarizing for Sixth Graders</td>
<td>(40)</td>
<td>Notes, References</td>
<td>trainer</td>
<td>09/14/04 05:10 PM</td>
</tr>
</tbody>
</table>

**noelle**

09/14/04 05:17 PM

**students’ summaries are too short**

I have been assigning a lot of summarizing activities to my class to help the students prepare for the upcoming tests. Many of my students have been having trouble writing long enough summaries - their summaries are too short! Does anyone have any suggestions that I could try in class next week? Thanks in advance for your help.
Recommended reading is automatically interjected into the discussion.

Using the mouse over feature, a user is able to select a note that looks interesting.

Summarizing helps students organize and evaluate information.

If you take away the value of them willing to die because no one cares if they do, you may have a chance to influence part of...
Assessment of Scenario Responses: Knowledge Post vs. Paper & Pencil

- Collected responses from over 200 officers at different posts
- Officers’ responses graded by two military experts
  - 72 TLAC responses (50% online, 50% paper)
  - 181 TKML responses (30% online, 70% paper)
- Higher quality responses using KP
- Demonstrable learning using KP
Air Force Academy

• Compared:
  – Instructor led classroom discussion
  – KP without expert interjections
  – KP with expert interjections
Expert Interjections

- Expert responses from LTCs were collected
- Expert interjections
  - Automatically added 1 expert comment for every 7 cadet comments
  - LSA selected expert comments based on current cadet discussion
US Air Force Academy trial

- Automated moderator in collaborative discussion over terrorist scenario
- 126 U.S. Air Force Cadets
- Significantly improved quality of discussions
  - Improved ideas generated by cadets
Predicting Team Performance scores based on team dialogue ($r = .76$)
Separating NATO Policy from Planning/Coordinating Comments

• Human classified over 1600 comments
  – Human classification correlated with LSA 0.84
  – Only 5% of comments were misclassified

• Multi-dimensional Visualization Tool

• View relationships between comments in LSA’s multi-dimensional space