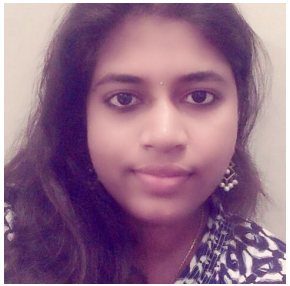
Tools for Detecting Exam Plagiarism WITHOUT Proctoring

Dr. Ed Gehringer, Guoyi Wang, Mounika Bachu
Department of Computer Science
North Carolina State University



http://tinyurl.com/iut-plagiarism-detection

Remote Proctoring

proctorU

Respondus

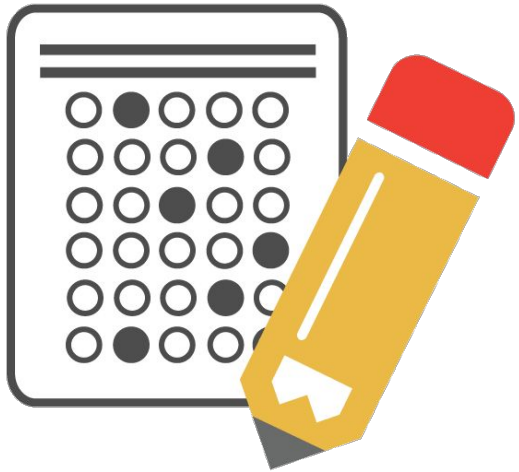
proctorio

KRYTERION

http://tinyurl.com/iut-plagiarism-detection

Plagiarism Tools for Multiple-Choice Questions

1. How they detect plagiarism
2. How to use the tools
3. How to interpret the results



http://tinyurl.com/iut-plagiarism-detection

How Multiple-Choice Plagiarism Detection Works

- Compare two students' answers
- A C B A D D A C B A 7 same answers
- A B B B D D A D B A of which 5 were wrong
- Of the 6 wrong answers that the first student had, 5 were shared with the second student.
- Metrics
  - How many answers were the same?
  - How many of those were wrong?
  - How many of the wrong answers were in common?
  - How frequently was each wrong answer chosen?

http://tinyurl.com/iut-plagiarism-detection

Three Plagiarism Tools

S-Check

CopyDetect

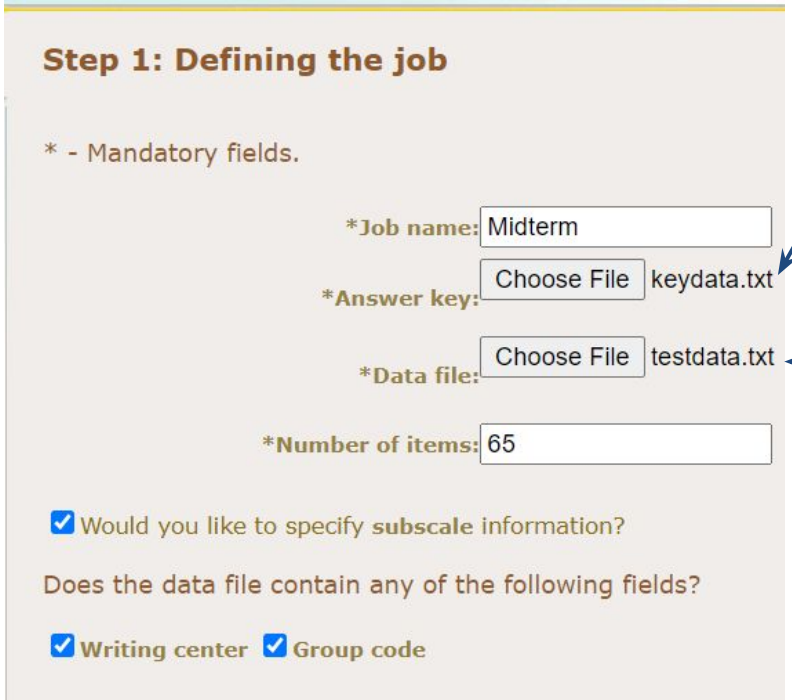
- Web app with a user-friendly interface
- Provides clear output reports which include detected pairs
- Requires a license

- Software created by George O. Wesolowsky (2000)
- Free, but time-limited
- Only contains one index
- Provides reports about detected pairs

- R package created by Cengiz Zopluoglu (2016)
- Free
- Contains more recently developed collusion indices
- No interpretation of output

http://tinyurl.com/iut-plagiarism-detection

INTEGRITY



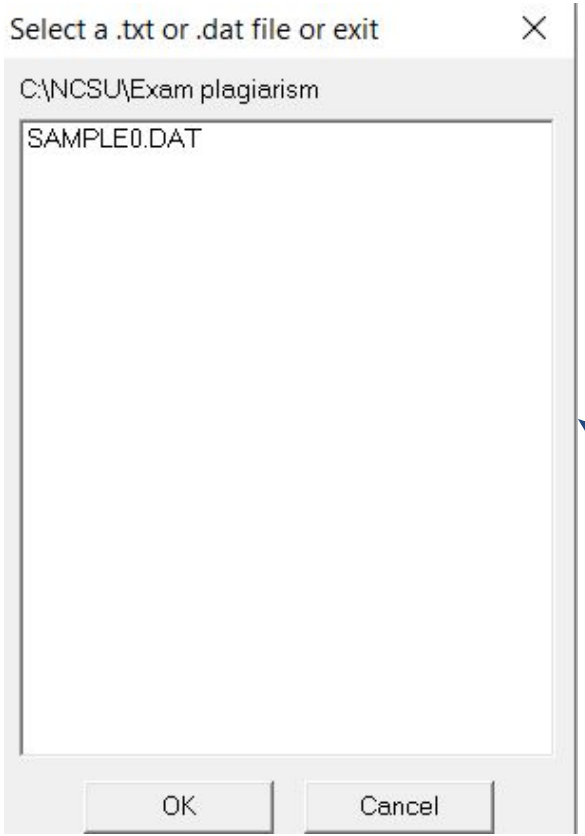
1,3,2,1,4,4,1,3,2,1

ID Class Group Responses

3,Class1,Male,3,1,4,3,3,2,2,4,1,4,1,2,4,2,4,2,3,1,1,3,3,3
6,Class1,Male,3,4,4,3,3,2,4,4,1,4,1,2,4,3,2,4,3,4,1,2,1,1
9,Class1,Male,3,4,4,3,3,2,3,1,4,3,1,2,1,3,4,3,2,4,1,3,3,3

http://tinyurl.com/iut-plagiarism-detection

S-Check



1. *Create a txt file*
2. *First line: key answers*
3. *Other lines: Responses*
4. *Change the extension (.txt) to .unm*

UNM file

```
answers  ,          ,          ,33143223141121331143132344244123244114244112224412
92933019 ,ABBEY     ,ERIC      ,33233244141121331143332444252421344221244123245132
52274024 ,ACHESON   ,ALEXANDE  ,32233244141122213124343214415212134411412414122 5332
84074012 ,BAXTER    ,MARIAN    ,33143244141121331143132444244123244114244122225412
```

ID	First Name	Last Name	Responses

UNMtoDAT.exe

DAT file

```
92933019,ABBEY,ERIC,..23..44...........3..4...524.13..221....23.4513.
52274024,ACHESON,ALEXANDE,.223..44.....21..2..4..1..152..13.....12..41..533.
84074012,BAXTER,MARIAN,......44..............4.................2...5...
```

1. *Correct responses are represented as dots*

CopyDetect

```
#R Script
install.packages("CopyDetect")
require(CopyDetect)
form1 <- read.csv("Form1.csv", header =
TRUE)
options<- c(1,2,3,4)
key_response <- c(2,3,1,4,1,2,2,1,1,1)
a <- similarity2(data = form1,
                 resp.options = options,
                 key = key_response,
                 person.id = "EID",
               center.id = “cent_id”,
                 item.loc = 3:12,
                 centers = c(19))
print(a)
```

	EID	cent_id	iresp.B1	iresp.B2	iresp.B3	iresp.B4	iresp.B5
49	e200049	19	4	3	3	3	2
52	e200052	19	1	2	1	2	1
103	e200103	19	2	3	2	3	1
120	e200120	19	2	3	1	4	1

Arguments	Description
data	A data frame, Columns: IDs, center IDs, Responses
resp.options	A vector of response options
key	A vector of key answers for all items
person.id	Column label for Student ID
center.id	Column label for Center ID
item.loc	The location of students’ responses
single.pair	A character vector for the suspected pair of examinees.
many.pairs	A matrix with two columns for students’ ID in one pair
centers	A character vector of center ids

Output from Plagiarism Tools

INTEGRITY

Detailed collusion detection report (all examinees)

	Examinee ID	Writing center	B-Index	PAIR1	PAIR2	MESA	g2
Pair 1	666 669	Class 3 Class 3	High 14.772	High 2535.000	High 3900.000	High 1.212E-017	High 12.324 12.324
Pair 2	672 675	Class 3 Class 3	High 9.536	Moderate 1170.000	High 1800.000	N/A	Moderate 6.742 6.742

Summary of Significances of Identified Pairs

pair	Z	A Priori Signif.	Bonferroni Signif.	Approx. Prog. Selected Signif.
9, 10	6.412	7.2E-11	1.8E-7	7.2E-11*
9, 12	5.349	4.4E-8	1.1E-4	1.1E-7
10, 12	6.412	7.2E-11	1.8E-7	7.2E-11*
18, 37	6.968	1.6E-12	4.1E-9	1.6E-12*
35, 68	4.535	2.9E-6	7.4E-3	4.9E-5
59, 71	3.919	4.4E-5	1.1E-1	2.6E-3

S-Check

CopyDetect

	Copier	Source	W.pvalu	GBT.pval	M4.pvalu	K.pvalu	K1.pvalu	K2.pvalu	S1.pvalu	S2.pvalu
76	e200477	e200372	0.009	0.026	0.012	0.098	0.001	0.015	0.02	0.256
106	e200372	e200477	0.009	0.026	0.012	0.098	0.001	0.015	0.02	0.256
92	e200120	e200420	0.027	0.032	0.037	0.273	0.237	0.304	0.183	0.35
99	e200511	e200420	0.027	0.032	0.019	0.122	0.193	0.134	0.164	0.335
56	e200049	e200266	0.894	0.954	0.962	1	1	1	0.974	0.813
80	e200103	e200406	0.898	0.897	0.929	1	1	1	0.99	0.846

Statistical Indices Used by Plagiarism Tools

Integrity S-Check CopyDetect

Statistical Distribution	Evidence of Answer Copying		
	Number of identical incorrect responses	Number of identical correct responses and incorrect responses	All items
Empirical Distribution	B (Angoff, 1972) Pair Ⅰ and Pair Ⅱ (Hanson, 1987)		
Normal Distribution		S-Check (Wesolowsky, 2000)	g_2 (Frary, 1977) ω (Wollack, 1997)
Binomial Distribution	K (Kling, 1979) K1 and K2 (Sotaridona, 2002) MESA		
Poisson Distribution	S1 (Sotaridona, 2003)	S1 (Sotaridona, 2003)	
Compound Binomial Distribution			GBT (van der Linden, 2006)

The important variables in plagiarism indexes

1. The evidence of answer copying
   a. Matching incorrect items only
   b. All matching items

2. Estimation of probability of a match
   a. Linear regression
   b. Item Response Theory

3. Statistical Distribution
   a. Empirical distribution
   b. Theoretical (Normal, Binomial, Poisson)

4. Similarity vs. copying
   a. Similarity indexes base on the performance level of both examinees. Each pair creates one index value.
   b. Copying indexes base on a suspected copier's performance level and the suspected source's answers. Each pair creates different index values for Copier-Source and Source-Copier

Empirical Distribution

The distribution of ***B***, ***Pair Ⅰ*** and ***Pair Ⅱ*** for all possible pairs of examinees will show the rarity of high indices values.

B : the number of identical incorrect responses from two students

Pair Ⅰ : the number of identical incorrect responses and length of the longest string of identical responses

Pair Ⅱ : the number of identical incorrect responses in the longest string of identical responses and the ratio of the number of identical incorrect responses to the sum of the non-matching responses and the number of identical incorrect responses.

Normal Distribution

$$Index = \frac{Obs - Exp}{SE}$$

g_2, **ω** indices investigates whether the observed number of matches exceeds it expected value. And they both need set one student as suspected copier and another one as a suspected source.

The probability, $P_C(u_{iS})$, is about someone with Copier's proficiency level selecting the Source's answer to item *i*. It is estimated based on **item difficulties**, **option difficulties**, the **score for Copier**, and the **particular answer selected by Source**.

g_2 uses a piecewise linear function to compute $P_C(u_{iS})$
ω uses the IRT model - Nominal Response Model (Bock, 1972).

S-Check: the probability, P_{jki} of a match (correct or incorrect responses)between students *j* and *k* on item *i* is computed based on **student ability** and the **proportion of students that answered the item correctly**.

Binomial Distribution

K, ***K*$_1$**, ***K*$_2$** use binomial distribution to compute the likelihood of matching on w_{ij} or more identical incorrect responses between two response vectors of a pair. K indices also need set one student as a suspected copier and another one as a source. ***K***, ***K*$_1$**, ***K*$_2$** use different way to estimating P_r

$$\sum_{u=w_{ij}}^{W_s} C\binom{W_s}{u} P_r^{\,u}(1-P_r)^{W_s-u}$$

w_{ij} : number of identical incorrect responses from student *i* and student *j*
W_s: number of incorrect items for source

$C\binom{W_s}{u}$: the number of all possible combination for u matches on W_s
r: group of copiers depends on the number incorrect responses.
P_r : the binomial probability of matching on an identical incorrect response with source and students in the *r*th incorrect-score which the copier belongs.

MESA is the index that **Integrity** modified from the traditional **ESA** method slightly to provide customized estimates of certain variables and to process information more quickly. **ESA** also uses a binomial distribution to model the number of identical incorrect responses between two students. It is similar to K indices, but it investigate the pair concurrently.

Poisson Distribution

***S*$_1$** uses a log-linear model to predict the number of identical incorrect responses between the suspected copier and source examinees and then uses a Poisson distribution to compute the likelihood of matching on wij or more identical incorrect responses between two response vectors of a pair.

$$S_1 = \sum_{u=w_{ij}}^{W_s} \frac{e^{-\hat{M}_r}(\hat{M}_r)^u}{u!}$$

w_{ij} : number of identical incorrect responses from student *i* and student *j*

$\hat{M}_r$: the model predicted number of identical incorrect responses with the source for the *r*th incorrect-score group.

W_s: number of incorrect items for source

r :group of copiers depends on the number incorrect responses.

***S*$_2$** is similar to ***S*$_1$**, but ***S*$_2$** includes number of identical incorrect and weighted identical correct responses.

Compound Binomial Distribution

GBT uses the generalized binomial distribution for the number of identical correct and incorrect response between two students. It is same as binomial, but probability of a match is allowed to be different for every item.

M4 (Maynes, 2014) uses the generalized trinomial distribution.

GBT and ***M4*** are both use the IRT model (Nominal Response Model) to estimate the probability of a matching answer.

A large body of research literature exists on how useful these methods are in practical settings.

GBT and **ω** are routinely found to be the most powerful.
(Wollack & Eckerly, 2017, Zopluoglu & Davenport, 2012, Zopluoglu, 2016)

K, K_1, K_2, S_1, S_2: acceptable, but pretty conservative, especially K-index
(Yormaz & Sunbul, 2017)

Plagiarism Tools for Essay Questions

1. Commercial software
2. LMS extension
3. Free software package
4. Rubric-based tests
5. Cloze test

http://tinyurl.com/iut-plagiarism-detection

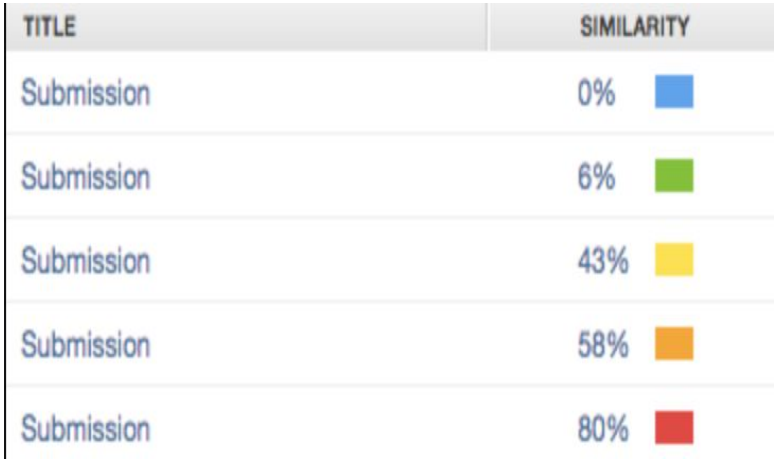
Turnitin

- An Internet-based plagiarism-detection service that compares submitted answers to several databases using a proprietary algorithm.
- Scans its own databases and also has licensing agreements with large academic proprietary databases which include internet, copyrighted pages from books, newspapers, and journals.

http://tinyurl.com/iut-plagiarism-detection

How to use Turnitin

- After logging in, click on All classes which allows you to add a new course or add new TAs or add new assignments.
- Click on the course and add the new assignment button and fill in the details
- Click “View”, and it displays the similarity percentage of all the students, right beside their name.



http://tinyurl.com/iut-plagiarism-detection

Plagiarism Detection Through an LMS

- **Unicheck** is a cloud-based plagiarism-detection application that finds similarities, citations and references in texts.
- Can be integrated into an LMS (Learning Management System) via a plugin, LTI, API or LTI+API
- Works with (at least) Canvas and Moodle
- To find similarities and paraphrases, checks are performed against the Internet, open-source repositories, and users' internal libraries or databases.

http://tinyurl.com/iut-plagiarism-detection

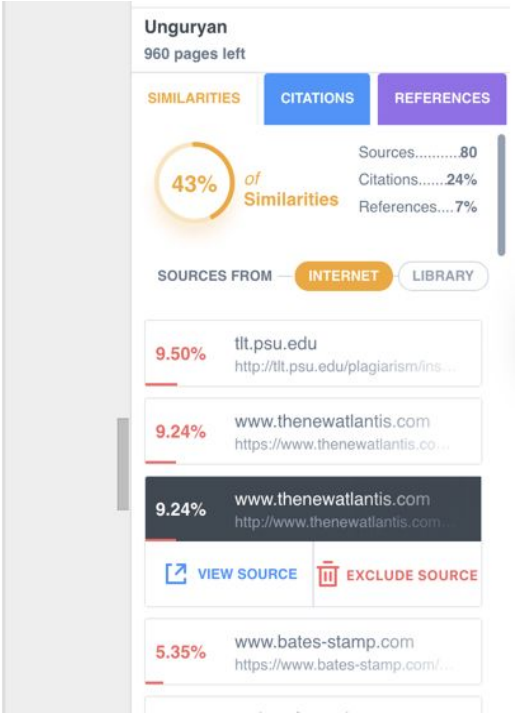
Unicheck Output

Fear the solution is sure to be very effective too. No doubt, that some may tend to find it very influential, since fear is one of the strongest emotions, especially when it is provoked by the threat of being expelled, e.g.:

Smith, Jones, and Parks note that "you can shorten a parenthetical note by naming the author of the source in the body of the essay; then the parenthetical note consists of a page number only" (782).

In his Little Book of Plagiarism, Richard Posner, a federal judge and prolific author, shows some understanding for the frustration and even the envy that might drive talented writers to such depths. He begins the book with an epigraph from the fourth-century Roman grammarian Aelius Donatus which expresses an annoyance every writer has felt: "Perish those who said our good things before we did." Posner applies the maxim to Viswanathan's misdeeds, writing that "in an age of specialization ... a creative person is apt to have a feeling of belatedness — a feeling that though just as creative as his predecessors he has appeared on the scene too late; the ship has sailed; the niche he might have filled has been filled already." A discerning reader might detect a note of sarcasm, however, as he continues: "Oh, the unfairness, Viswanathan might have thought, of McCafferty's having picked the low-hanging 'chick-lit' fruit rather than leaving some of it for her." Posner's judicious mix of sympathy for the motive with harsh criticism of the infraction colors his book throughout.

Studies that examine links between cardiovascular and mental activity must understand that cardiovascular activity itself comprises a suite of variables (Van Roon, Mulder, Althaus, and Mulder, 2004).



http://tinyurl.com/iut-plagiarism-detection

Textreuse

- An R package which provides a set of functions for measuring similarity among documents and detecting passages which have been reused.
- Useful in cases like detecting duplicate documents in a corpus prior to text analysis, or for identifying borrowed passages between texts.

How to use it:

1. Install R package
2. TextReuseCorpus(paths, dir = NULL, text = NULL, meta = list(),
   progress = interactive(), tokenizer = tokenize_ngrams, ...,
   hash_func = hash_string, minhash_func = NULL, keep_tokens = FALSE, keep_text = TRUE, skip_short = TRUE)

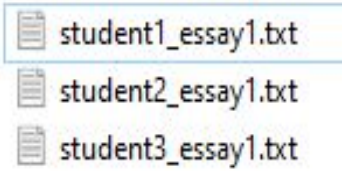
http://tinyurl.com/iut-plagiarism-detection

Arguments	
paths	A character vector of paths to files to be opened.
dir	The path to a directory of text files.
text	A character vector (possibly named) of documents.
meta	A list with named elements for the metadata associated with this corpus.
progress	Display a progress bar while loading files.
tokenizer	A function to split the text into tokens
…	

Example of code

a) dir<-"path of the directory"

i) The directory will include txt files which are the responses of each student and each essay question.



ii) corpus <- TextReuseCorpus(dir = dir, tokenizer = tokenize_ngrams, n = 1)

a) tokenize_ngrams and n=1 mean the sentence will splitted by one word.

iii) comparisons <- pairwise_compare(corpus, jaccard_similarity)

1.jaccard_similarity

2.jaccard_dissimilarity

3.Jaccard_bag_similarity

4.ratio_of_matches

iv)Comparisons

1) Similarity

	student1_essay1	student2_essay1	student3_essay1
student1_essay1	NA	0.015625	0.02884615
student2_essay1	NA	NA	0.02884615
student3_essay1	NA	NA	NA

2) Dissimilarity

	student1_essay1	student2_essay1	student3_essay1
student1_essay1	NA	0.9919355	0.9873418
student2_essay1	NA	NA	0.9870968
student3_essay1	NA	NA	NA

3) Ratio of matches

	student1_essay1	student2_essay1	student3_essay1
student1_essay1	NA	0.125	0.1271186
student2_essay1	NA	NA	0.1261261
student3_essay1	NA	NA	NA

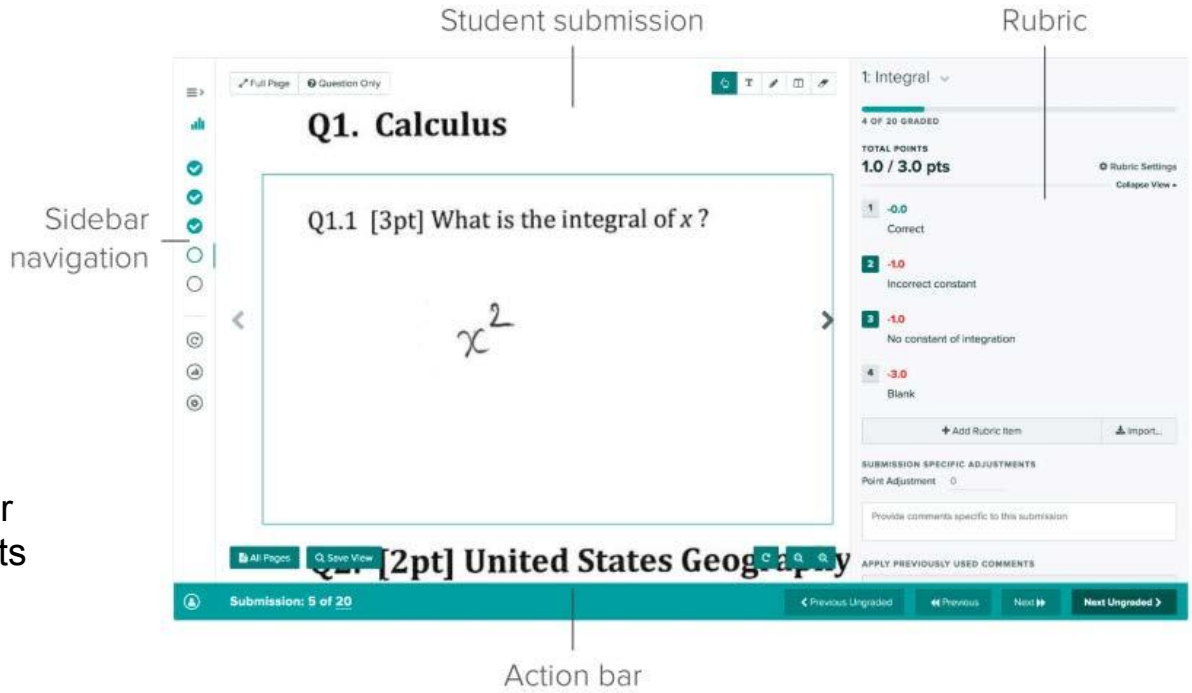
http://tinyurl.com/iut-plagiarism-detection

Rubric-Based Tests

gradescope

From Gradescope, now a unit of Turnitin.

Compares rubric scores for different students, highlights similarities.



http://tinyurl.com/iut-plagiarism-detection

Cloze Test

- Students with suspicious submissions can be required to take this test.
- They are asked to fill in the words that have been removed from their submission

Example:

Why is education so important in our life?

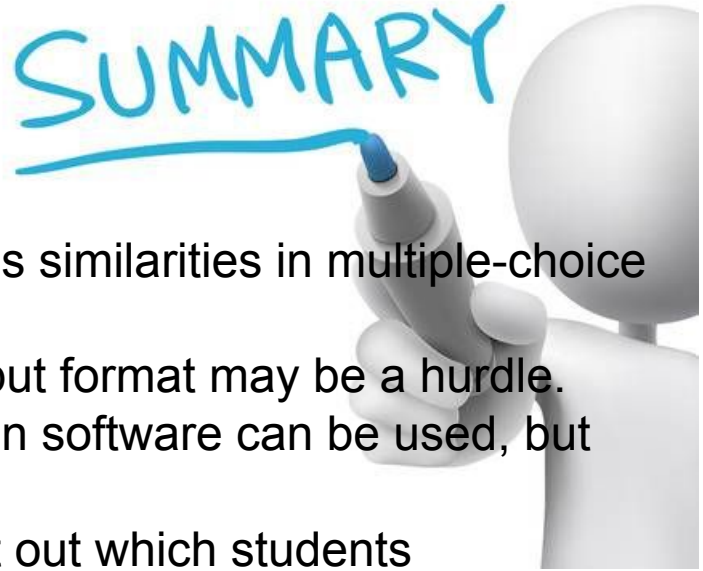
Student's answer: Education gives us knowledge of the world around us and changes it into something better. It develops in us a perspective of looking at life. It helps us build opinions and have points of view on things in life. People debate over the subject of whether education is the only thing that gives knowledge.

Chase procedure:

Education gives us knowledge of the _____ around us and changes it into something _____. It develops in us a perspective of _____ at life. It helps us build opinions ___ have points of view on things in _____. People debate over the subject of _____ education is the only thing that gives _____.

- The person who has stolen the answers will not be able to do it and can be sanctioned.

http://tinyurl.com/iut-plagiarism-detection



- Statistical tests can warn of suspicious similarities in multiple-choice exams
- A variety of tools facilitate this, but input format may be a hurdle.
- For essay exams, plagiarism-detection software can be used, but again, input format may be a hurdle.
- Cloze tests can be employed to ferret out which students understand suspicious exam answers.
- Gradescope's rubric-based test is a promising approach.

http://tinyurl.com/iut-plagiarism-detection

Thank you