

Short-Answer Questions [25 points]:

1. Investigating one search engine (Google or Bing), and figure out what tricks they are using. (15%)

(a) Does it use stemming? How do you know? (Explain your testing method in all the following sub-questions as well)

(b) Does it filter stop words?

(c) Does it use spell correction?

[ANS]

I investigated Google search engine for whether it uses stemming, filters stop words or uses spell correction. The sentence used for examination is “*in a exciting store at Frogs of the nightmares*”, a part of the result is as below:

Stuck in nightmare sequence. :: Balrum General Discussions

<https://steamcommunity.com/app/424250/discussions/0/412449508294160504/> ▾

Mar 15, 2016 - **Store** Page. View Stats: Global Achievements · Balrum > General Discussions > Topic Details. Regalis Ferus · View Profile View Posts. Mar 15, 2016 @ 3:13am. Stuck in **nightmare** sequence. The dream after talking to the **frog**. I'm in an empty **nightmare** with no way out. Pretty lame: <. 1, 1. > Showing 1-2 of ... Missing: exciting

Frog's breath | Etsy

https://www.etsy.com/market/frog's_breath ▾

Shop for **frog's** breath on Etsy, the place to express your creativity through the buying and selling of handmade and vintage goods.

Missing: exciting

A Semi-Definitive List of Worst Nightmares by Krystal Sutherland ...

<https://www.barnesandnoble.com/w/a-semi-definitive...nightmares-/1125320219> ▾

Sep 5, 2017 - The Hardcover of the A Semi-Definitive List of Worst **Nightmares** by Krystal Sutherland at Barnes & Noble. FREE Shipping on \$25 or more! ... Hearts, was released in 2016, and was published in over twenty countries. Her three greatest fears are heights, dark caves ... and (perhaps worst of all) **frogs**.

[PDF] The Frog Prince Sorry, Wrong Number The Actor's Nightmare

www.lexhamarts.org/theater/200504/program200504.pdf ▾

The **Frog** Prince. The **Frog** Prince by Sally Bailey with music and lyrics by Carol Gulley. Sorry, Wrong Number by Lucille Fletcher. The Actor's **Nightmare** by Christopher Durang. He would also like to thank David for another opportunity to work with a wonderful director; he's so **excited**, he's willing to lose his head for the ...

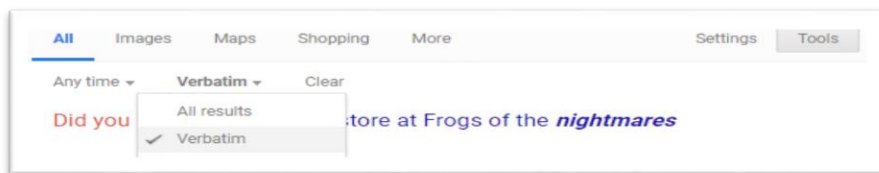
First, Google uses stemming, which is an approximate method for grouping words with a similar basic meaning together. For example, the word set “frog” and “frogs” generally have the same meaning, and so as the word set “nightmare” and “nightmares”. These words are all marked in bold style by Google search engine in the result. Obviously, it is an evidence that Google is using stemming method for querying.

Second, Google has filtered stop words. In my query, stop words like “in”, “at”, “of” and “the” is included. While in the returning result, none of these words are marked in bold style even some of them appear in result documents. As a result, we believe that google has filtered these words during querying.

Third, Google is using spell correction for querying, yet it provides option for user to not to use spell correction. For instance, I made a typo “nightmars”. In response, google corrected it for me, returned result documents base on the correct sentence and provided me options if I want to query without auto correction.

Showing results for in a exciting store at Frogs of the *nightmares*
Search instead for in a exciting store at Frogs of the nightmares

On the other hand, Google search engine has an option called “Verbatim”, which removes the functions we discussed above. According to [Google Introduces Verbatim Searching](#) written by *Rob D. Young*, while using “Verbatim” option, the search engine will only use the same tense, the same verb form and the same plural/singular form for querying. As a result, only “frogs” will be targeted, only “nightmares” will be targeted, and no defined stop words anymore.



Un No KODOT: comics search

www.ohnorobot.com/index.php?self=1&page=1&s=plane&comic=48 ▼

[At the OLD HANG TREE< on a tiny planet.]] Dr Quickly: Oh. Cool. Dr Quickly: And THIS must be a SPACE CRYSTAL. Some quest! Just a trip to "the corner store", really. Interstellar Frog: WHO GRABS UP THE CRYSTAL OF THE FROGS OF THE SPACE WHO DAAAAARES Dr Quickly: Type: Dr Quickly: Well, the crystal is ...

not in bold style

A Semi-Definitive List of Worst Nightmares by Krystal Sutherland ...

<https://www.barnesandnoble.com/w/a-semi...of...nightmares...sutherland/1125320219> ▼

Sep 5, 2017 - So are haircuts, spiders, dolls, mirrors and three dozen other phobias she keeps a record of in her semi-definitive list of worst nightmares. Then Esther is pickpocketed ... His patterned button up—clearly donated to a thrift store sometime in the mid- 1970s—was torn at the collar. The boy sobbed

2. For a particular search query, your IR system returns 14 relevant documents and 16 irrelevant documents. There are a total of 80 relevant documents in the collection. (10%)

(a) What is the precision of the system on this search?

[ANS] By definition, precision is the fraction of relevant documents among the result documents. Therefore, the precision is $14 / (14+16) = 46.7\%$

(b) What is the recall of the system on this search?

[ANS] By definition, recall is the fraction of relevant document that have been retrieved over the total amount of relevant documents. Therefore, the recall is $14 / 80 = 17.5\%$