

# Googling to the Max - Exercises

Getting the most from: <http://www.google.com>

Google Toolbars (FREE) - highlight terms, remember searches, search within a site, block pop-ups, and more:

For Internet Explorer & Firefox

<http://toolbar.google.com>

For other Mozilla browsers

<http://googlebar.mozdev.org>

## #1 HOW GOOGLE "THINKS" IN DEFAULT MODE CRAFTING BASIC GOOGLE SEARCHES

AND automatically implied between terms

- all your terms somewhere  
in text of pages  
in pages that link to a result page  
in other pages on the same site

Ex. 1: What is matched on?

1. Search Google for the keywords: **google page rank**
2. Click Cached: for the page titled "Google Technology" or "Google PageRank Calculator"
3. What is the explanation of the matching of your terms with these pages?

STEMS some words

- finds word with various endings  
search **kite flying** and get matches on **kite, kites, kiting and flying, fly, flies**
- turn off with + or " " as in **+kite +flying** or **"kite flying"**

IGNORES common or "stop" words

- when this happens, a gray message appears below the search box in results telling you what was ignored
- turn off with + or " " just as for stemming

RANKING FAVORS pages with your words in PHRASES, CLOSE TOGETHER, and IN THE ORDER TYPED

Ex. 2: Word order and word choice matter

Compare the top results for searches in Google for these three sets of keywords:

<b>grass snake</b>	Mostly about a kind of snake. By stemming also matches <b>snakes</b> and <b>grasses</b>
<b>snake grass</b>	Many about a kind of grass, not present in 1 <sup>st</sup> search
<b>snake in the grass</b>	Most pages contain this expression, even though Google said <i>in</i> and <i>the</i> were ignored; also finds pages with <b>grass snake</b> and a few with <b>snake grass</b>
<b>snake +in +the grass</b>	Eliminates pages not containing <i>in</i> and <i>the</i> somewhere.
<b>"snake in the grass"</b>	Most specific and precise. Requires exact phrase in all pages.

OR searching requires capitalized OR

- Can be used between single words and phrases enclosed in quotes:  
**california OR oregon OR "pacific coast"**  
**"global warming" OR "greenhouse effect"**

No parentheses for nesting; keep OR searches simple

## #2 EXPLOITING GOOGLE "FUZZY" SEARCH OPTIONS

### □ PUNCTUATION THAT IS NOT IGNORED

- Apostrophe ( ' ) : peoples, people's, and peoples' are searched as different words
- Hyphen ( - ) : same-sex retrieves same-sex, same sex and samesex  
Always supply the - to search any word that might be used hyphenated
- Accent marks in Roman-alphabet foreign languages : éléphant does not match elephant (and vice versa)  
If searching in the language where the accent is common, the accents are not required for matches. Google assumes people writing web pages in that language may or may not put the accents in.
- Asterisk ( \* ) : Whole word wild card, allows one or more unspecified words.

### □ SYNONYM SEARCHES ( ~ )

- Google will "think" of words with similar meaning: ~food matches recipes, nutrition, cooking  
~facts matches information, statistics  
~help matches guide, tutorial, FAQ, manual

### □ SIMILAR PAGES (in results list) or command related:[URL]

- Google will "think" of pages like the one you choose, by using links to and from the page, words in the page, and the importance of pages in links.
- Uses: evaluate a questionable page by the links in and out  
find comparable pages when shopping or looking for a type of site  
broaden a search without thinking of words that might be in pages like the one you like

#### Ex. 3: Exploiting "FUZZY" Google options

1. Find pages about: *The one-child law of the People's Republic of China*, knowing that some people never use apostrophes properly:  
one-child law people's OR peoples "republic of china"
2. Search a name allowing varying intervening words using \*:  
george \* \* bush
3. To find out, *How do praying mantises hear?* try searching:  
~ears praying mantis
4. Find pages similar in focus to another page:  
Search for consumer webwatch  
At the end of the website [www.consumerwebwatch.org](http://www.consumerwebwatch.org), click [Similar pages](#).

### □ NUMBER RANGE allowed in searches ( .. without spaces)

- Follow search terms with beginning and ending numbers, separated by two periods.
- Can be used one sided, as *less than* or *greater than*

#### Ex. 4: Number range searches

1. Find pages mentioning Babe Ruth between 1921 and 1935  
"babe ruth" 1921..1935
2. Find digital cameras priced below \$300 with resolution of 4.5 to 5 megapixels (MP)  
digital camera ..\$300 4.5..5 megapixels OR MP

### #3 SETTING LIMITS IN GOOGLE SEARCHES

- Require search terms to occur in specific parts of web pages
  - **intitle:** Words must occur in the official <title> field in the head part of the page
  - **site:** Limit to a site or domain (first part of the URL, before the first / )
  - **inurl:** Require terms to occur anywhere in URL (URL punctuation ignored)
    - . Can be combined with each other and with other search terms

Ex. 5: Limiting to parts of web pages

1. Find pages primarily about mileage in hybrid cars:  
`intitle:mileage "hybrid cars"`
2. Find pages about President Bush and either global warming or greenhouse effect:  
`bush intitle:"global warming" OR intitle:"greenhouse effect"`
3. Is there anything on this subject from the official Whitehouse site?  
`site:www.whitehouse.gov "global warming" OR "greenhouse effect"`
4. Find pages containing sheet music written to commemorate the assassination of Abraham Lincoln  
`site:memory.loc.gov lincoln "sheet music"`  
Searches within all Library of Congress's American Memory Project pages that are in Google.  
`inurl:lincoln "sheet music"`  
Finds pages with "sheet music" in them and with "lincoln" somewhere in the URL.

- Limit to type of document or file extension: **filetype:**
  - Google has many types of files besides HTML pages (doc, xls, wpd, ppt and more)
  - View as HTML link in results with these file types lets you see the document without the application and avoids any virus or worm they could contain

Ex. 6: Limiting to type of document or file:

1. Find fact sheets and reports (not just web pages) about the death-rate for malignant mesothelioma in the U.S.  
`filetype:pdf death-rate malignant mesothelioma "united states"`
2. Find PowerPoint presentations from reputable universities in the U.S. about the impact in China of its one-child policy  
`filetype:ppt site:edu china one-child policy`

- Google Advanced Search does not allow *intitle:*, *inurl:*, or more than one *site:*, and has few date range options.
- Pages from within a foreign country by building on the URL of a Google search result
  - Uses information about where the webpage server is located, as well as country coding; more comprehensive than what is available a [http://www.google.com/language\\_tools](http://www.google.com/language_tools)
  - Search on the terms you want. In your browser's Address or URL box, append to the result (without spaces):  
`&restrict=countryXX` with XX being the CAPITALIZED two-letter country code  
To find country codes: Search Google for TLD [country name] or  
List of all country codes available at [http://en.wikipedia.org/wiki/List\\_of\\_Internet\\_TLDs](http://en.wikipedia.org/wiki/List_of_Internet_TLDs)

Ex. 7: Pages from within a country:

- Find recipes for authentic Moroccan cuisine, from within Morocco
- Step 1: perform the search:  
`recipes OR cuisine morocco OR moroccan`
- Step 2: append to the resulting URL in the Address box `&restrict=countryMA`  
`http://www.google.com/search?q=recipes+OR+cuisine+morocco+OR+moroccan&restrict=countryMA`

## #4 HANDY GOOGLE TOOLS AND SHORTCUTS

- DEFINITIONS, THESAURUS, and ENCYCLOPEDIA LOOK-UPS
  - `define:[your words]` Finds web-based definitions, Google-ranked, and often reliable
  - `search and click your word in statistics bar` Access to published dictionary entries and other resources through Answers.com. Options vary for different words.

Ex. 8: Definitions and more:
1. Definitions: <code>define:internet</code> <code>define:due diligence</code>
2. Locate dictionary definitions and thesaurus terms (synonyms, antonyms, etc.) using the link in the blue statistics bar after you search: <code>scholarship</code> <code>internet</code>

- TRANSLATIONS
  - Click [Translate the page](#) after a page in a foreign language
  - Paste a URL or a piece of text at [http://www.google.com/language\\_tools](http://www.google.com/language_tools)

Ex. 9: Translating:
1. Go to Google Advanced Search - <a href="http://www.google.com/advanced_search">www.google.com/advanced_search</a>
2. Limit to PAGES IN FRENCH and search <code>jacques chirac avec george bush</code>
3. Click <a href="#">Translate this page</a> (appears with most pages in languages not in your search interface)
4. You can also translate by pasting a URL or text - <a href="http://www.google.com/language_tools">www.google.com/language_tools</a>

- CALCULATOR and CONVERSIONS - parentheses accepted to sequence and clarify expressions
  - Mathematical functions (+, -, \*, /, % of, *n*th root of, sqrt(*nn*), ^ for exponentiation)
  - Advanced math (Trig: sin, cos, tan, sec, csc, cot, etc.; Inverse trig: arcsin, arcos, arctan, etc.; Hyperbolic trig: sinh, cosh, tanh, etc.; Logarithms; Exponential functions; Factorials; and more)
  - Many mathematical constants (pi, imaginary numbers, and more)
  - Units of measure and conversions
    - . MASS: kilograms or kg, grams or g, grains, pounds or lbs, carats, stones, tons, tonnes, etc.
    - . LENGTH: meters or m, miles, feet, inches, Angstroms, cubits, furlongs, etc.
    - . VOLUME: gallons, liters or l, bushels, teaspoons, pints, drops, etc.
    - . AREA: square miles, square kilometers, acres, hectares, etc.
    - . TIME: day, seconds or s, years, centuries, sidereal years, fortnights, etc.
    - . ELECTRICITY: volts, amps, ohms, henrys, etc.
    - . ENERGY: Calories, British thermal units or BTU, joules, ergs, foot-pounds, etc.
    - . POWER: watt, kilowatts, horsepower or hp, etc.
    - . INFORMATION: bits, bytes, kbytes, etc.
    - . QUANTITY: dozen, baker's dozen, percent, gross, great gross, score, etc.
    - . NUMBERING SYSTEMS: decimal, hexadecimal or hex, binary, roman numerals, etc.
    - . QUICK FACTS: currency in countries, population, biographies
  - Physical constants (atomic mass units or amu, Avogadro's number, Boltzmann constant, Faraday constant, gravitational constant, mass of a proton, mass of each planet and of the sun, permeability of free space, etc.)

Ex. 10: Calculating and converting:
1. How many carats in a pound? 1 lb in carats
2. How many square feet in 2.5 hectares? 2.5 hectares in square feet
3. Value of 1555 in hexadecimal? 1555 in hex
4. The mass of an electron? mass of an electron
5. How many bits in a 2.5 gigabytes? 2.5 gigabytes in bits
6. What is the currency in Uganda? currency in uganda

- SHORTCUTS to many other kinds of publicly available information
  - Area codes (U.S.), defined by maps - type 3 digit code
  - Airport conditions - type airport code or city and the word airport
  - Flight tracking - type the word airline and the flight number
  - Maps - type an address
  - Stock info - type one or more NYSE, AMEX, NASDAQ, or mutual fund codes
  - White pages - type rphonebook: name address
  - Yellow pages - type bphonebook: name address
  - Zip codes - type a zip code
  - Movies - type movies: city, st or movies: zip code
- STANDARD NUMBER LOOKUPS
  - Parcel tracking for UPS, FedEx, USPS - type the number
  - US patent lookup - type the word patent and the number
  - UPC product identification - type the code
  - VIN vehicle info, history - type the number

Ex. 11: Shortcuts and standard number lookups:
1. What part of the U.S. is in the zip code 87455 87455 Look below the map at the town and state.
2. Are there currently any delays at O'Hare airport in Chicago? chicago airport or ord airport
3. What's playing in Berkeley? movies:berkeley, ca

- UNIVERSITY SEARCHES
  - Conduct a search within the website of a university - you specify the topic
  - <http://www.google.com/options/universities.html> has long list of University links to start searching from

Ex. 12: University searches:
Go to Google Advanced Search. At the bottom, click on <u>Universities</u> . Choose any university from the list and search for information on: graduate admissions

## #5 THE BEST OF GOOGLE'S FAMILY OF DATABASES

- Images
  - One of the largest images database - 1.3+ billion (Yahoo Images boasts 1.5 billion)
- News
  - 4,500 news feeds, 30 days of news
- Shopping
  - Froogle - shopping sites from Google web page database + merchant-supplied machine-readable catalogs
- Local (and Maps)
  - Draggable images, satellite view, local (yellow pages), driving directions
  - Google web pages + several yellow page directories
  - For "finding businesses and services, two search boxes: enter a place or address and a topic
- Groups
  - Google Groups - Usenet Newsgroups back to 1981
- Videos
  - Videos submitted by people, from broadcasts and TV
  - Viewable in Flash if ► button in the thumbnail
  - Search close captioning to locate
- More in Google Labs - labs.google.com
  - Google Sets - generate items in a set you start
  - Google Suggest – as you type, suggests options for search terms (also in Japanese)

Ex. 13: Sampling Google databases:

At the Google homepage (<http://www.google.com>), click on more>>

1. Go to [video.google.com](http://video.google.com) and find a film called "What old people do for fun."
2. Go to Advanced Image Search. Find full color pictures in JPG format of mona lisa.
3. Go to [local.google.com](http://local.google.com)

Drag the US map so that an area of California is centered in the screen

Drag the map to reposition it in your hometown (or another area)

Click Satellite (upper right)

Click Local and search for something like restaurants or hotels

Try driving directions between two addresses

## #6 GOOGLE BOOKS SEARCH and GOOGLE SCHOLAR

Two of Google's initiatives directly involve published information and access to published information

Google Scholar ([scholar.google.com](http://scholar.google.com))

- Search for journal articles in scholarly publications
- Large, interdisciplinary database - from citations in web pages and from some journal publishers
- Advanced search allows author and other searches
- Integrated with UC-elinks

Google Books Search ([print.google.com](http://print.google.com))

- Full-text of some books with links to where to buy
- Promises to digitize millions of publications in libraries - may link to library holdings if public domain

Ex. 14:

1. In [print.google.com](http://print.google.com), search rock climbing peru
2. At [scholar.google.com](http://scholar.google.com), search for an article by somebody Koontz or Coontz about how the institution of marriage has changed, been transformed, and reflects transformations in society over the ages.

In Scholar Advanced Search; marriage required in TITLE and koontz OR coontz in AUTHOR

## Table Comparing **GOOGLE BOOKS** and **GOOGLE SCHOLAR**

Two of Google's initiatives rather directly involve libraries, published information, and access to published information, both on web pages and in print.

	<b>Google Books</b>	<b>Google Scholar</b>
<b>Google's goal</b>	To "organize the world's information and make it universally accessible and useful" focused on content of books	To "search specifically for scholarly literature" in many formats
<b>How to use</b>	Through the special search box at <a href="http://books.google.com">books.google.com</a>	Through special a search box at <a href="http://scholar.google.com">scholar.google.com</a>
<b>What's included</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Publisher-supplied publications; also authors of self-published and out-of-print publications can submit their works as the publisher</li> <li><input type="checkbox"/> Publications supplied by selected libraries (Harvard, Stanford, U of Michigan, Oxford, NYPL), most public domain</li> <li><input type="checkbox"/> Google scans, digitizes, and indexes words in each book for free (no charge, no profit)</li> <li><input type="checkbox"/> Mostly books, but some articles (primarily <i>Variety</i>, <i>Library Journal</i>, <i>School LJ</i>, <i>Publishers Weekly</i>, and technical, scientific, and pharmaceutical publications)</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Mostly articles, but some citations to books</li> <li><input type="checkbox"/> Full-text of online articles found by Google's robot-like crawlers from academic publishers, professional societies, preprint repositories, and scholarly articles posted on the web</li> <li><input type="checkbox"/> Citations to articles and books found in lists of references in publications found by Google's robot-like crawlers</li> </ul>
<b>What you get</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Scanned full text to search and browse</li> <li><input type="checkbox"/> Printing of pages or images if public domain</li> <li><input type="checkbox"/> Snippets or excerpts of copyright books when publisher has not permitted full text online</li> <li><input type="checkbox"/> Copy of table of contents and index</li> <li><input type="checkbox"/> Link to publisher, bookstores, and libraries if a scanned library copy</li> <li><input type="checkbox"/> Paid advertisements related to your search, as in Google web search</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Full text if freely available or if provided by the publisher with permission</li> <li><input type="checkbox"/> Ucelinks available when we have a subscription (other universities too – set in Preferences)</li> <li><input type="checkbox"/> Works with Proxy Servers</li> <li><input type="checkbox"/> Citation or abstract (not full text) if protected by copyright</li> <li><input type="checkbox"/> Links to citations to most articles (like a "who cites whom")</li> <li><input type="checkbox"/> Occasionally, "Library Search" tries to locate a library near you that owns the book. Partial implementation of OCLC holdings links</li> </ul>
<b>Search tips</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Search from <b>books.google.com</b></li> <li><input type="checkbox"/> <b>Advanced Search</b> <a href="http://books.google.com/advanced_book_search">http://books.google.com/advanced_book_search</a></li> <li><input type="checkbox"/> <b>intitle:</b> to limit</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Go to <b>scholar.google.com</b> or <b>Scholar Advanced Search</b> <a href="http://scholar.google.com/advanced_scholar_search">http://scholar.google.com/advanced_scholar_search</a></li> <li><input type="checkbox"/> Enter author name, publication name, and/or subject words</li> </ul>
<b>More information</b>	<b>About Google Books</b> <a href="http://print.google.com/googleprint/about.html">http://print.google.com/googleprint/about.html</a>	<b>About Google Scholar</b> <a href="http://scholar.google.com/scholar/about.html">http://scholar.google.com/scholar/about.html</a>