Biol 381 – Bioinformatics Lab
Lab 1: Database Searching and Literature Analysis

Introduction

Much of research involves finding out what is already known about a topic you'd like to study. This means that it's important to develop skills that allow you to get the information you need, as well as digest that information. In today's lab session, we will focus on two ways to learn new information about a topic that are beneficial for everyone interested in any kind of research to do well: searching the web, and reading and analyzing primary scientific literature.

The first topic we will address is how to effectively search NCBI's resources for information about a topic. Most people these days have used Google to search the internet, and in many ways, searching NCBI is very similar. But not everyone is aware of the strategies that can be used to improve the effectiveness of their searches, such as Boolean logic (see Figure 1), parentheses and quotation marks, and knowing different terms that refer to the same subject. These tools can help you find relevant information quickly by broadening or narrowing your search. Today, you'll have an opportunity to practice using them and improve your search skills.
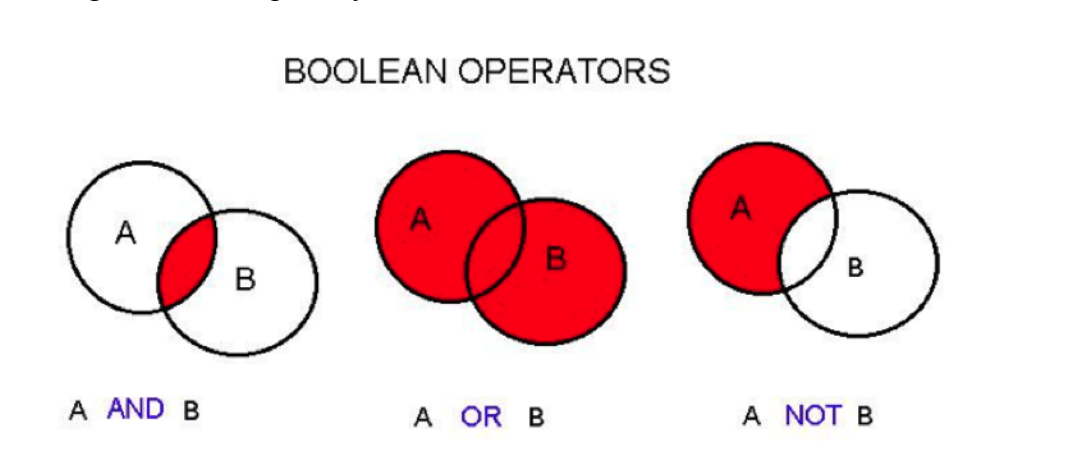

Figure 1: Scopes of Boolean operators. Adapted from http://libwebdev2006.fullerton.edu/images/ librarycontent/Boolean.jpg.

In addition to making sure your search is effective, it's important to analyze your results to determine which of them are likely to be helpful and relevant to the topic you want to study. For any search you do, it's probable that some of your results will be *false positives*—hits that are not related to your search topic. Hopefully, many of your results are *true positives,* meaning that they are actually relevant to your topic. The positives—the results returned by your search—are all well and good, but it's highly likely that there are relevant resources out there that didn't appear in the results—*false negatives*. Finally, all the other articles in the database that didn't come up in your search, but that also aren't relevant, are *true negatives*.

While this may seem like a strange way to think about search results, it is a concept that will come up

throughout this course, so you should begin to familiarize yourself with the basics. You may find Figure 2 helpful as a way to visualize the differences between true and false positives and negatives. For a more in-depth explanation of the concept, you can also refer to Figure 2.13 on p. 41 in your textbook.
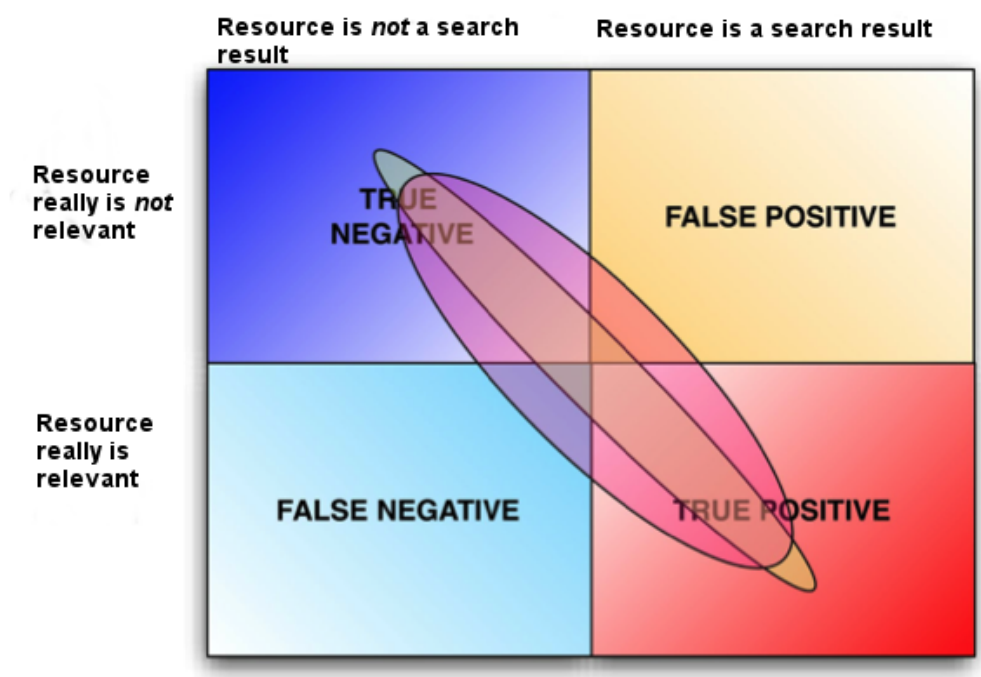

Figure 2: True and false positives and negatives, as they apply to internet search results. Adapted from http://100dialysis.files.wordpress.com/2012/05/falsetrue.jpg.

The second topic that we will focus on today, and the one that is often trickier to learn, is how to read and analyze scientific literature. Primary research papers are often dense, and may contain terms and techniques that are unfamiliar to you. However, they are also often extremely valuable references when it comes to scientific topics, so being able to understand their content is very useful.

When you read scientific literature, the way that you approach an article makes a difference. While it might seem tempting to read a paper from beginning to end like a book, this strategy might not be the best one if you want to make sure you know what's going on in the study. Many guides to reading scientific literature point out that it might make more sense to read the sections of the paper in a different order—it's a good idea to read a few of these guides to see what they suggest and why. Additionally, it might be a good idea to take notes as you read, noting the answers to the questions at the end of this handout, and keeping track of terms you're unfamiliar with and their definitions. If the researchers used a technique you haven't seen before, it's probably a good idea to look it up.

One final note: if you get stuck, don't be afraid to use resources other than the ones we've provided for help! NCBI has lots of tutorials and help pages on searching and navigation, and you may also find your textbook helpful. Of course, there's always Google as well.

Part 1: NCBI Searching

In the first part of this exercise, you'll get a feel for how to use NCBI's resources to obtain information about a topic. NCBI has many databases associated with it, and analyzing results from all of them would take a much larger amount of time than we have for this exercise. For this reason, we will focus on one database today, NCBI Bookshelf.

NCBI Bookshelf is a database of free textbooks, manuals, and other forms of literature related to biomedical science. All of the Bookshelf resources are full-text and searchable.

Step 1:
Before you begin searching NCBI to learn about your topic, you should make sure you have a good understanding of the strategies you can use to make your search as effective as possible. To familiarize yourself with how NCBI searching works and how to view your results, go to the NCBI Bookshelf page and look at some of the tutorials and help pages (there should be links at the bottom of the page). If you find an especially helpful article or video, share it with your group!

Step 2:
Each person in your group should design an NCBI query to find articles that relate to your topic. Remember to use the search strategies we discussed at the beginning of this lab to maximize the number of results that are actually related to your topic ("true positives"). Keep refining your query until you're happy with the quality of your results.

Step 3:
Look at the first 20 results of your NCBI search. Read the titles, and try to determine whether each result will actually be helpful in answering your question. (If the title alone doesn't make this clear, or even if you just want further confirmation, you can also skim the abstract or introduction to get a better idea of what the article or book is about.)

- How many results on the first page are true positives? How many are false positives?

Step 4:
Share the query you designed with your group, as well as the number of true positives in your first 20 search results.

As a group, discuss the effectiveness of each query.

- Which query was the most effective in terms of true:false positive ratio? Why do you think that query worked well? Be specific, and give as many reasons as you can.

- How do you think the other queries could be altered to improve their effectiveness?

Part 2: Reading and Analyzing Scientific Literature

In the second part of this exercise, you'll focus on analyzing scientific literature to learn more about your topic.

Step 1:
Your group will be given a paper to read and analyze. Before you begin reading the paper, review some of the strategies we have discussed, as well as the guides we have provided, as they will provide a good basis for making sure you really understand what you read.

Here are some additional tips that you may find helpful in your analysis:

- Remember to keep track of unfamiliar terms, and look them up if necessary. It's a good idea to take notes, either in the article's margins or on a separate sheet of paper. You may also want to keep track of any abbreviations that appear in the paper and what they stand for, in order to remind yourself as you read.

- Look over the questions on the last page of this handout, and keep them in mind as you read. You will discuss the answers to these questions later with your group.

- If you come across a procedure that is unfamiliar to you, you can look it up using NCBI Bookshelf. In particular, *Molecular Biology of the Cell* by Alberts is a resource that can help you understand experimental techniques. Feel free to use other resources as well.

- Even though you may have studied some of the procedures in the past in other courses, it's still a good idea to refresh your memory about the techniques discussed in your paper. You'll want to know what each technique used for, as well as what the steps are in the process.

In addition to discussing the topic of your paper, you may also want to discuss the following questions with your group:

- Do you think it makes sense to read a paper from beginning to end, or is there a better way to read that will help you make sure you know what's really going on?

- In what order did you read the sections of this paper? Keep track of the order, and share it with your group in your discussion of the paper.

Step 3:
Discuss the paper as a group. Go over the answers to the handout questions, and make sure that everyone understands what they read. Additionally, discuss the answers to the questions below.

- Did you come across any concepts, techniques, results, or conclusions that were difficult to understand? What were they? What helped you understand them better?

- In what order do you think this paper should be read? Why?

Step 4:
As a group, present your paper. Each person in your group should choose a figure from the paper to explain to the class. Be sure to explain what is represented by the figure, as well as what those results mean in terms of the study.

Reading Questions

1. Title and First Author:

2. What is the purpose of the study?

3. What is the hypothesis?

4. What were the results of the study?

5. Give a brief explanation of how the study was done.

6. What was previously known about the topic of the paper?

7. What conclusions do the authors reach?

8. What are the most significant results of the study, and why are they significant?

References

1. Bogucka, R., & Wood, E. (2009). How to read scientific research articles: a hands-on classroom exercise. *Issues Sci Technol Libr*.