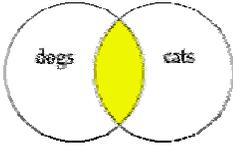
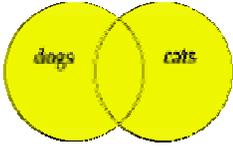
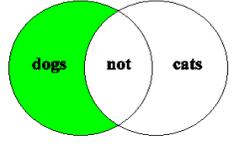


Boolean Searching for the Web

Boolean searching is named after George Boole, a British mathematician (1815-1864), who wrote about logical ways to formulate precise queries using true-false connectors or "operators" between concepts. The true-false nature of Boolean logic, as this system is commonly called, makes it compatible with binary logic used in digital computers. It has become the conventional basis for searching most computerized systems. The primary operators are **and**, **or**, and **and not**. **near** is a modification of AND. You join words or concepts with these operators. Parentheses are used to organize the sequence and groups of concepts.

OPERATOR	WHAT IT DOES & WHEN TO USE IT
<p>and</p>  <p>dogs and cats is the small overlap</p>	<ul style="list-style-type: none"> • Requires all terms (words or phrases-in-quotes) joined by it to appear somewhere in the document, in any order. • Using and narrows a search, because it augments the number of matching words required for a document to meet the criteria of the search. • The more words you enter connected by and, the fewer documents you get. • USES: <ul style="list-style-type: none"> ○ The and operator is generally used to join different kinds of concepts, different aspects of the question. ○ "global warming" and "sea level rise" and california - All three terms (<i>global warming, sea level rise, and california</i>) must be somewhere in every document retrieved.
<p>or</p>  <p>dogs or cats includes any page with at least one of the terms</p>	<ul style="list-style-type: none"> • Requires at least one of the terms joined by it to appear somewhere in the document, in any order. • Using or broadens a search, because it makes it easier for documents to meet the criteria of the search: any of the terms will suffice. • The more words you enter connected by or, the more documents you get. • USES: <ul style="list-style-type: none"> ○ The or operator is generally used to join similar, equivalent, or synonymous concepts. You would be happy to find any one of them in the pages you retrieve. ○ "global warming" or "greenhouse effect" - You will accept either (or both phrases) in retrieved pages. ○ california or oregon or alaska or "pacific coast" - At least one must be in each retrieved document.
<p>and not</p>  <p>dogs and not cats excludes pages that mention cats, even if they also mention dogs</p>	<ul style="list-style-type: none"> • Excludes documents containing whatever follows it. • Using and not limits a search, because it disqualifies documents no matter whether they meet the other criteria of the search or not. • The and not operator is generally used after you have performed a search, looked at the results, and determined that you do not want to see pages containing some word or phrase. • USES: <ul style="list-style-type: none"> ○ The and not operator should be used with extreme caution, because it eliminates the entire page, and some pages may be of value to you for other information they contain. I almost never use and not for this reason. ○ "global warming" and "sea level rise" and not california - The first two terms must be somewhere and any page containing <i>california</i> will be thrown out.
<p>near, w/s, w/p, w/#, and other proximity search operators</p> <p>dogs near cats requires both terms, like and, with the added requirement that they be within 10 words of each other</p>	<ul style="list-style-type: none"> • Requires the term following it to occur within a certain proximity of the preceding word in the search. In AltaVista Adv. Search, near requires the terms to be <i>within 10 words</i> of each other in either direction. There are many variations on "near." • Joining words by near gives you fewer documents than and, because it requires the words be closer together. • USES: <ul style="list-style-type: none"> ○ The near operator is used when you want to require that certain terms appear in the same sentence or paragraph of the document. ○ "global warming" near "sea level rise" - Requires the two phrases to occur within 10 words of each other, in either direction, in all pages retrieved.

<p>() parentheses: "Nesting"</p>	<ul style="list-style-type: none"> • Require the terms and operations that occur inside them to be searched first. When more than one element is in parentheses, the sequence is left to right. This is called "nesting." • Parentheses MUST BE USED to group terms joined by or when there is any other operator in the search. <ul style="list-style-type: none"> ◦ "global warming" and "sea level rise" and (california or "pacific coast*") - Requires first two terms somewhere in all documents, and either <i>california</i> or <i>pacific coast</i>. • Parentheses MUST BE USED with near to specify which terms must occur within 10 words of each other when there is any other operator in the search. <ul style="list-style-type: none"> ◦ ("global warming" near "sea level rise") and (california or "pacific coast*") - Requires <i>sea level rise</i> to be within 10 words of <i>global warming</i>; the rest can be anywhere in the pages. The parentheses guarantee that the effect of near stops with <i>sea level rise</i>. • Keep nesting simple!! It is possible to make very complicated, multiply nested Boolean searches, with parentheses within parentheses. In web searching, that is <i>not good practice</i>. Web searching is free, and doing several simpler searches is usually better than trying to create the perfect, humongous search. With simpler searching, you have a good sense what to alter to improve results; with a huge search, it's hard to tell what part retrieved what page.
<p>Implied Boolean searching, or "pseudo-Boolean"</p>	<ul style="list-style-type: none"> • Most search engines assume the and operator between words. This is what is called "default AND search." • When - is used to exclude, it is the equivalent of the and not operator. • There are different ways in which search engines may implement the or operator. It is not always available.

You do not need or even want to get very complicated with Boolean searching in web searching. Searching is Web is free, and several simpler searches take less time than a humongous search. Moreover, with complicated searches, you often don't know which parts of the search worked and which did not. Simpler searches can more easily be compared with one another, and you know what worked.

How much Boolean logic do I have to learn to use?

The **minimum** everyone needs to search databases effectively:

- You want to feel comfortable with **or** searching in at least two of the Advanced search engines. It is particularly important in AltaVista Advanced because **or** in the Boolean expressions box gives you 100% flexibility in combining words, phrases, and limiting of any type (much more than any other search engine or even AltaVista Advanced's boxes). Each of the search engines requires a different technique to achieve **or** searching. (They each think their non-standardized variation on **or** searching is the easiest for you to learn.)
- You must remember to put parentheses () around terms connected by **or** or **near** in AltaVista Advanced. I do not recommend more advanced nesting of terms within terms. In fact, Google and AllTheWeb do not allow you to make complicated, nesting searches.