

Alchemist v2.0:

A GUI-based tool for analyzing morphemes and creating morphological gold-standards in XML format

General Purpose of Alchemist

The original purpose of Alchemist is to allow you to read in raw text files and create morphological gold-standards in XML format. Using Alchemist, you can identify morphemes, along with a number of important characteristics of the morphemes, such as whether they are roots or affixes, the degree of analyst certainty, and allomorphs of the morpheme.

Alchemist is also a good general tool for sorting and filtering lists of words, because it allows the user to easily use regular expressions applied to words.

User Documentation Contents

[Key Terms and Features](#)

[Basic Operation](#)

[Reading in a New Text File](#)

[Using Alchemist - The Basics](#)

[Alchemist User Interface](#)

[Advanced Features](#)

[Word Scrubbing](#)

[Scrubbing Data](#)

[Tips for Using Regular Expressions](#)

[Common Regular Expression Symbols](#)

[Using Regular Expression Variables](#)

[More Resources for Creating Regular Expressions](#)

[Sorting Data](#)

[Filtering Data](#)

[Regular Expression Filter Examples](#)

[Sample Filter Search](#)

[Word Collection Table](#)

[Assigning Morpheme Status](#)

[Deleting Words from the Word Collection Table](#)

[Viewing Changes in the Word Collection Table](#)

[Morpheme Explorer](#)

[Merging Morphemes](#)

[Show Filtered](#)

[Clearing Cuts](#)

[Word/Morpheme Information](#)

[Morpheme Attributes](#)

[General Word/Morpheme Information](#)

[Saving Data](#)

[Guide to Navigation Tabs and Menus](#)

Key Terms and Features

<i>Feature</i>	<i>Definition</i>
scrub	This feature allows you to clean up text files before analyzing words in the Word Collection table. There are simple and advanced scrubbing options in the Word Scrubbing window.
sort	This feature allows you to organize how you would like to view your data, either alphabetically or reverse alphabetically (see Sorting Data section).
filter	This feature allows you to create regular expression searches to narrow down the group of words you analyze (see Filtering Data section).
cut	This feature refers to morpheme assignments. You can <i>make cuts</i> , such as assign New Root or New Affix status to morphemes. You can also <i>clear cuts</i> , which means you can remove previously assigned morpheme status.
merge	This feature, which is located in the Morpheme Explorer , allows you to group together morphemes and allomorphs.
Word Collection table	The table that displays all the words from the corpus you feed into Alchemist to analyze.
Morpheme Explorer	The window where you organize and view the information you assign to morphemes.

Basic Operation

Reading in a New Text File

To import a new text file into Alchemist, go to the File menu and select **New** (shortcut Ctrl+N). This will prompt you to enter a pre-defined number of distinct words that you would like to analyze in your collection. The default number of distinct words is 500.

Next, you will be prompted with a Word Scrubbing window, which allows you a variety of beginner and advanced options for manipulating the text that will appear in the Word Collection table. For the most basic use, click **Scrub** to accept simple scrubbing options and begin editing in Alchemist. To ignore the scrubbing options and read the text file in as its original form, click **Skip**. For more details on how to *scrub* the text you will be working with, see the below section on [Word Scrubbing](#).

Using Alchemist - The Basics

1. Once your text appears in the Word Collection table, you can label morphemes occurring within words by using the mouse to highlight sequences of letters in one or more words. To better organize your data, use the Sorting or Filters features to align whole sets of words with similar roots or affixes (see sections on [Sorting Data](#) and [Filtering Data](#) for more guidance on using these features).
2. Once you have selected one or more occurrences of a morpheme, assign it a status (e.g., *R* for *New Root* or *A* for *New Affix*). These *cuts* will immediately appear in the Morpheme Explorer window just below the Word Collection table. Root assignments

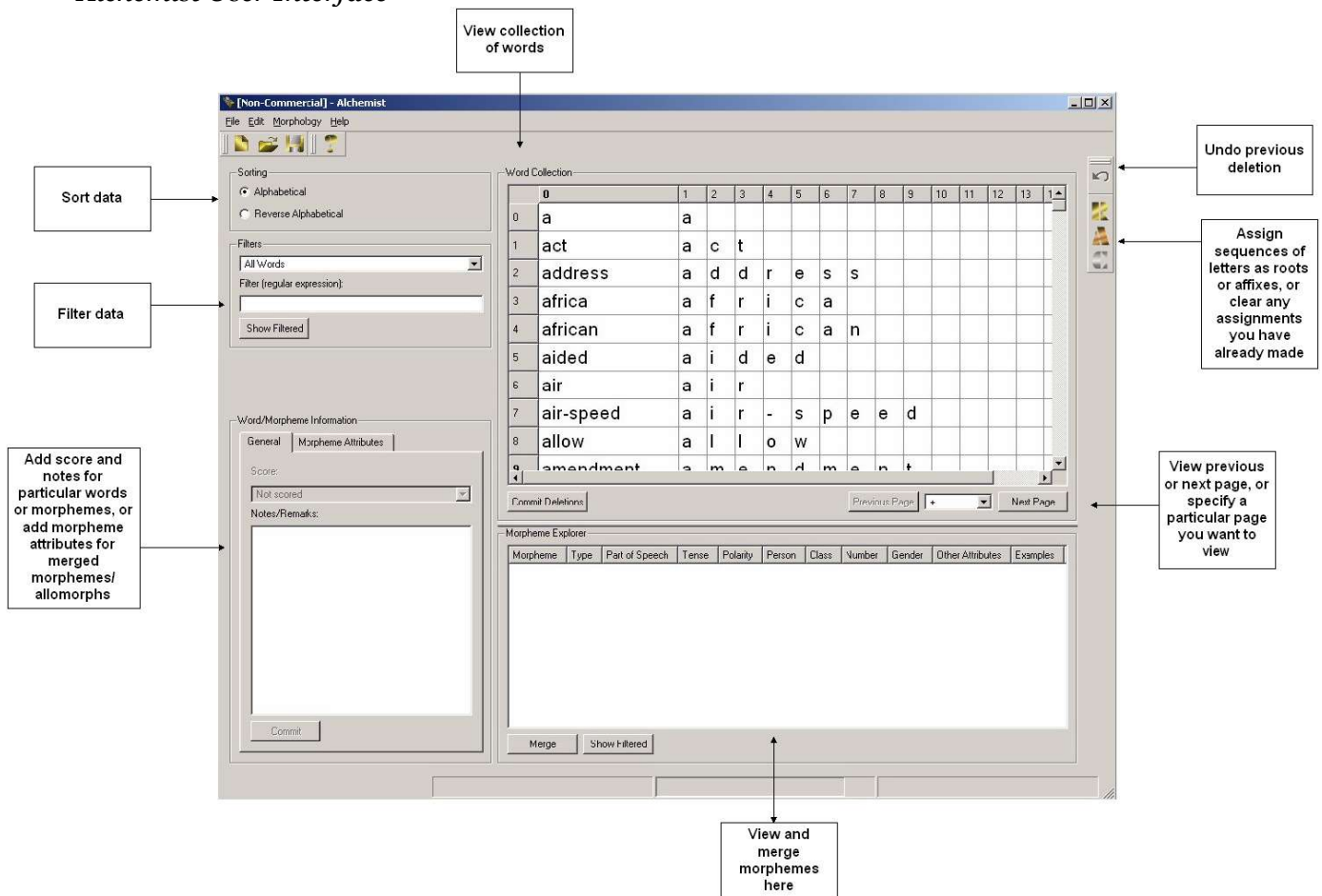
will be gold and silver, affix assignments will be bronze and copper, and undefined substrings will remain white.

3. You may also enter general information for a particular morpheme or set of morphemes by highlighting the applicable morphemes and marking your certainty about the analysis as well as adding specific notes or remarks. Click **Commit** to save any notes or remarks that you add.
4. To merge all the selected morphemes or allomorphs in the Morpheme Explorer into one entry, highlight the applicable morphemes and click the **Merge** button.
5. Next, you will be prompted to enter Morpheme Attributes (e.g., morpheme type, tense, person, number, part of speech, polarity, class, and gender) in the Word/Morpheme Information section, located in lower left-hand corner of the application. This information will be automatically saved as you enter it.
6. To save the collection to a pre-chosen XML (*.xml) file, go to the File menu and click **Save** (shortcut Ctrl+S). This XML file can be reopened and edited in Alchemist.

Alchemist User Interface

The following diagram shows the high-level features of the Alchemist application.

Alchemist User Interface



Advanced Features

Word Scrubbing

Scrubbing Data

You may choose either simple or advanced scrubbing options. If you prefer not to alter the text you feed into Alchemist, click **Skip** to begin using Alchemist.

Under the **Scrubbing** tab, you can set your simple preferences for how you want to clean up your text files before analyzing words in the Word Collection table. For example, you can lower-case all the words in your file and remove any numbers. In addition, you can remove preceding, internal, and following punctuation. The text entry boxes that allow you to specify punctuation are activated by default and pre-populated with the most common punctuation that may be useful to scrub; however, you can add, edit, delete, or de-activate any of the information based on your scrubbing needs.

Basic Scrubbing Options

The screenshot shows a dialog box titled "Scrub files" with a "Load" button and a "Save" button. Below these are two tabs: "Scrubbing" (selected) and "Advanced". The "Scrubbing" tab contains four sections, each with a checked checkbox and a text input field:

- ☒ Make lower case
- ☒ Remove numbers
- ☒ Remove preceding punctuation:
Input field: `"'${[({<`
- ☒ Remove internal punctuation:
Input field: `.,`
- ☒ Remove following punctuation:
Input field: `.,?!:"'"}>/>`

At the bottom of the dialog are three buttons: "Help", "Scrub", and "Skip".

Under the **Advanced** tab, you can create and customize new regular expression replacement rules for scrubbing data. This means you can replace every occurrence of the **Find (regular expression)** with the **Replace with** string. Once you add a rule, it will appear at the bottom of the pre-populated ordered list of replacement rules from the **Scrubbing** tab (e.g., Remove numbers, Remove preceding punctuation, Remove internal punctuation, and Remove following punctuation). This list contains all the scrubbing rules that will be used, organized in the order in which they will be applied, from top to bottom. All rules can be re-ordered (using the **Move up** and **Move Down** buttons) and removed (using the **Remove** button) by clicking on the buttons to the right of the list. You may also add, edit, or delete the existing list of preferences.

Advanced Scrubbing Options

Scrub files

Load Save

Scrubbing Advanced

New replacement

Find (regular expression):

Replace with:

Add

Make lower case
Remove numbers
Remove preceding punctuation
Remove internal punctuation
Remove following punctuation

Move up
Move down
Remove

Help Scrub Skip

Once you have entered your preferences, click **Scrub** to update your text file and begin editing in Alchemist.

You may also save your preferred set of scrubbing options to a file that you can re-load in Alchemist for later analysis or load for use in another program. To add pre-existing scrubbing options, click the **Load** button and browse to the file location. To save new or updated scrubbing preferences, click **Save**.

Tips for Using Regular Expressions

Using regular expressions to scrub data allows you to efficiently and systematically create replacement rules that modify how your text appears when you analyze it in the Word Collection table. To add a new replacement rule, click on the **Advanced** tab and enter a regular expression for the strings you want Alchemist to both find and replace.



NOTE

Regular expression searches are *case-specific*; therefore, if you scrub your data to be all lowercase, you must create regular expressions that also use lowercase characters.

The simplest form of regular expression is a character (e.g. 'a' or 'b'). An expression can also be a set of characters. For example, using brackets to enclose characters (e.g., '[ABCD]' or as shorthand, '[A-D]') will match to any character in the set (e.g., 'A' or 'B' or 'C' or 'D'). For example, the search '[A-D]' will return all words in your collection that include any of these capital letters (e.g., 'Albany', 'Birmingham', 'Chicago', 'Detroit').

Common Regular Expression Symbols

The following table gives some examples of the regular expression symbols that you may choose to add to strings in either your Advanced scrubbing preferences or when filtering data in the Word Collection table.

<i>Symbol</i>	<i>Definition</i>	<i>Example</i>
.	(dot) any single character (i.e., word character, whitespace character, punctuation, digit)	b.d → body, Cambodia b..d→ bird, borderline
\s	whitespace character only	These options, while similar to the above symbol (.), allow you to filter more narrowly, i.e., if you want to search for characters that specifically include or exclude whitespace, word characters, punctuation, or digits.
\S	non-whitespace character (i.e., word character, punctuation, digit)	
\w	word character only (i.e., NO whitespace, punctuation, digit)	
\W	non-word character (i.e., whitespace, punctuation, digit)	
\d	digit only	
\D	non-digit	
\n	n th back-reference	\1, \2 → reference sets of characters to be found and replaced (see section on Using Regular Expression Variables)
\$	end of word boundary	ing\$ → searches for words that end with the string '-ing'

Symbol	Definition	Example
^	beginning of word boundary or used to negate the character set if it occurs as the first character, i.e. immediately after the opening square bracket	^a, ^av → searches for all words starting with the character 'a-' or the string 'av-', respectively [abc] matches 'a' or 'b' or 'c', but [^abc] matches anything except 'a' or 'b' or 'c'
\b	word boundary, i.e., any non-word character such as a space, a new line, or the beginning or end of a string	Similar to above examples for \$ and ^, except that it does not specify boundary location
*	Kleene star: matches to an occurrence of 0 (zero) or more times	^ad* → searches for words that begin with a- and have 0 (zero) or more occurrences of d following it; i.e., this brings back all words beginning with 'a-', since 'd' may occur 0 (zero) times
+	Kleene plus: matches to an occurrence of 1 or more times	^ad+ → searches for words that begin with 'a-' and have 1 or more occurrences of a d following it; this brings back only words with the sequence 'ad-', since 'd' must occur at least once
{#,#}	sets a range where a character must occur at least as many times as the first numeral, and at most as many times as the second numeral	b{1,2} → searches for words where the character 'b' occurs at least once and at most twice in a row; if you do not want to specify a minimum or maximum, leave it blank (e.g., {,2} or {1,}, respectively)
?	used to make an expression optional	ds?\$ → searches for words that end in 'd' and optionally 's'
	used for alternation (i.e., OR expression)	(av ad) → searches for words including the string <i>av</i> OR <i>ad</i>
[]	used to enclose sets of characters	^[a-d] → searches for words that begin with any letter ranging from 'a' to 'd' (i.e., 'a', 'b', 'c', 'd')
()	used to group expressions together as a unit and to identify parts of the regular expression that you wish to capture. This is useful for building up more complex regular expressions (see section on Using Regular Expression Variables).	^(av ad) → searches for words that begin with 'av-' OR 'ad-' (e.g., 'available', 'address')



TIP

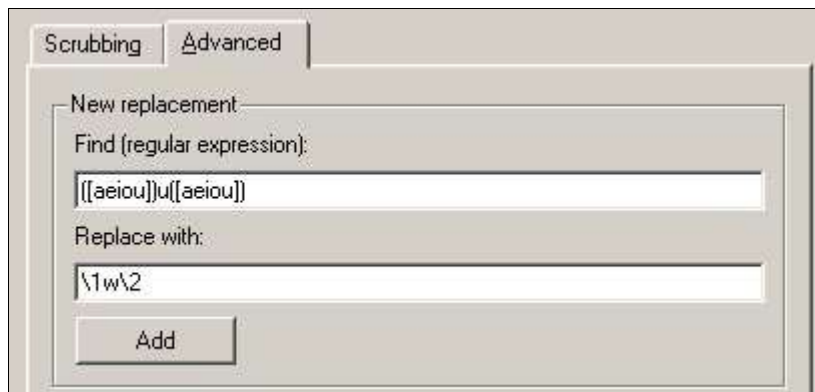
Do not use spaces when writing regular expressions unless you intend for the search to include the whitespace. For example, the above example `^(av|ad)` would not return the same results had you written it `^(av | ad)`.

Using Regular Expression Variables

When creating regular expression replacement rules, be sure to isolate only the part of the string that you want to be changed. To avoid deleting the parts of a regular expression that you want to remain unaffected, use variables to match to sets of parentheses. Variables are marked as numbers delimited with the escape character '\', which correspond to a particular parenthetical set. Following is an example of a replacement rule using variables.

The following replacement rule searches for every occurrence of the letter 'u' that occurs between two other vowels with a 'w'. The resulting replacement for the word 'naua' will be 'nawa'. If the two parenthesized expressions were not replaced as variables, the vowels on either side would be lost and the result would be 'nw'. As shown below, '\1' refers to the first set of parenthesis '([aeiou])' and '\2' refers to the second set '([aeiou])'.

New Replacement Rule in Advanced Scrubbing Window



More Resources for Creating Regular Expressions

For additional references and tutorials on regular expressions, visit the following websites:

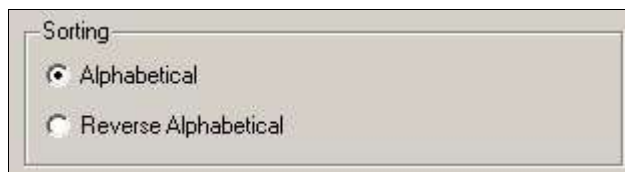
Regular-Expressions.info: <http://www.regular-expressions.info>

Qt Reference Documentation: <http://doc.trolltech.com/3.3/qregexp.html>

Sorting Data

Data can be sorted in two ways, either from left-to-right or from right to left. The Sorting feature is located in the upper left-hand side of the application:

Sorting Options



Alphabetical – This is the usual left-to-right way that words are alphabetized (i.e., all words beginning in the same sequence of letters come together in the list), in ascending order beginning with numerals and the letter 'A'. This option is useful for identifying prefixes and similar roots with different suffixes.

Reverse Alphabetical - This option sorts data from right-to-left (i.e., words are aligned at the right-most column), so that all words ending in the same sequence of letters come together in the list. The words are displayed in ascending order, beginning with numerals and the letter 'A'. This option is useful for identifying word endings, such as suffixes, suffix combinations, and similar roots with different prefixes.

Filtering Data

The Filters feature is located just below the Sorting feature. To filter data:

1. Choose one of three filters:



All Words – Searches all words in the file.

Analyzed Only – Searches only words you have made affix or root assignments to.

Unanalyzed Only – Searches only words you have not made affix or root assignments to.

2. Enter your filtering search into the **Filter (regular expression)** text entry box.



Regular Expression Filter Examples

Below are some regular expression examples that may be useful when filtering data:

Regular Expression	Function
<code>(\w{2,})\1</code>	Finds reduplications of length 2 or more with no intervening characters.
<code>(\w{2,})\w*\1</code>	Finds reduplications of length 2 or more which may have intervening characters.
<code>a\w*e\w*i</code>	Find words containing 'a', 'e', and 'i', in that order, possibly with intervening characters.
<code>ing</code>	Find words containing the string 'ing' anywhere in the word.
<code>ing\$</code>	Find words containing the string 'ing' at the end of the word only.

For more information on how to find or write your own regular expressions, see the section on [Tips for Using Regular Expressions](#).

Sample Filter Search

You can filter data by limiting the sequence of letters you want searched for in the collection. For example, if you want all words that contain a particular suffix string such as '-ing', then enter that string in the **Filter (regular expression)** box. You are then presented with only the words containing that string, and you can more easily identify the words among them that contain that particular morpheme.

Filter (regular expression): ing

Sorting		Word Collection										
<input checked="" type="radio"/> Alphabetical <input type="radio"/> Reverse Alphabetical		0	1	2	3	4	5	6	7	8	9	10
Filters		0	hosting		h	o	s	t	i	n	g	
All Words		1	huntington		h	u	n	t	i	n	g	t
Filter (regular expression):		2	meaning		m	e	a	n	i	n	g	
ing		3	painting	p	a	i	n	t	i	n	g	
Show Filtered		4	ring				r	i	n	g		

To further narrow the results so that you can easily identify your intended sequence(s), you may also choose to include or exclude particular text strings. For example, after determining that the string 'ring' never occurs in your text as part of a word with the suffix '-ing', you may choose to narrow your filtered set of words to include the string 'ing', but exclude the string 'ring'.

Filter (regular expression): `[^r]ing`

Sorting		Word Collection										
<input checked="" type="radio"/> Alphabetical <input type="radio"/> Reverse Alphabetical		0	1	2	3	4	5	6	7	8	9	10
Filters		0	hosting		h	o	s	t	i	n	g	
All Words		1	hun	tington		h	u	n	t	i	n	g
Filter (regular expression):		2	mean	ing		m	e	a	n	i	n	g
[^r]ing		3	paint	ing		p	a	i	n	t	i	n
Show Filtered		4	think	ing		t	h	i	n	k	i	n

In addition, you may want to exclude any words in which the string 'ing' occurs somewhere other than at the end of the word, such as in 'hungtingon'.

Filter (regular expression): `[^r]ing$`

Sorting		Word Collection										
<input checked="" type="radio"/> Alphabetical <input type="radio"/> Reverse Alphabetical		0	1	2	3	4	5	6	7	8	9	10
Filters		0	hosting		h	o	s	t	i	n	g	
All Words		1	mean	ing		m	e	a	n	i	n	g
Filter (regular expression):		2	paint	ing		p	a	i	n	t	i	n
[^r]ing\$		3	think	ing		t	h	i	n	k	i	n
Show Filtered		4	warn	ing		w	a	r	n	i	n	g

3. Click the **Show Filtered** button to view results in the Word Collection table.

Word Collection Table

All the words in the Word Collection table will be visible after opening a document. If filters have been applied, only a filtered subset of the collection will be visible.

Assigning Morpheme Status

You can select words or parts of words (i.e., morphemes) by clicking on a cell and dragging the selection area to encompass other areas. Entire rows or columns can also be selected by clicking on the row or column number at the left and top of the table.

Once you have selected one or more occurrences of a morpheme, you can assign it a status (e.g., *R* for *New Root*, *A* for *New Affix*, *C* to *Clear Cuts*). Right-clicking on a selected area brings up a context menu of commands that can be applied to the area. Alternately, you can click on the buttons located to the right of the Word Collection table. These cuts will immediately appear in the [Morpheme Explorer](#) window just below the Word Collection table. Root assignments will be gold and silver, affix assignments will be bronze and copper, and undefined substrings will be white.

Assigning Morpheme Status

	0	1	2	3	4	5	6	7	8	9	10	11
0	aided					a	i	d	e	d		
1	annotated	a	n	n	o	t	a	t	e	d		
2	called				c	a	l	l	e	d		
3	measured		m	e	a	s	u	r	e	d		
4	populated	p	o	p	u	l	a	t	e			
5	produced		p	r	o	d	u	c	e			
6	united				u	n	i	t	e			
7	used						u	s	e			

Context Menu:

- New Root Ctrl+R
- New Affix Ctrl+A
- Clear Cuts Ctrl+C
- Refresh F5
- Delete Del
- Set score... ▶

Buttons: Commit Deletions, Previous Page, aided (dropdown), Next Page

Deleting Words from the Word Collection Table

To delete a word from your collection, highlight the row by clicking on the horizontal row number of the word. If all cells of a word are selected, the entire word will be deleted from the collection; if not, only the part of the word that is selected will be removed.

Highlighting a Row for Deletion

	0	1	2	3	4	5	6	7	8	9	10
0	+	+									
1	a	a									
2	act	a	c	t							

Click the **Delete** key on your keyboard, or go to the Edit menu and select Delete.



TIP

If you delete a word from your collection, you must commit your deletion by clicking the **Commit Deletions** button. You will not be able to make any further edits until you do so.

Morpheme Explorer

Each time you create a morpheme in the collection, it is also made visible in the Morpheme Explorer. Morphemes can then be merged. This designates them as the same morpheme or as allomorphs, and they can then have common attributes assigned to them. The Morpheme Explorer displays the actual words that you cut in the Word Collection table to be included in the morpheme group (located under the heading Examples). You can also use the Morpheme Explorer window to filter the morphemes or sets of morphemes you view in the Word Collection table.

Morpheme Explorer Window

Morpheme Explorer											
Morpheme	Type	Part of Speech	Tense	Polarity	Person	Class	Number	Gender	Other Attributes	Examples	
ing	Affix									boiling	
ing	Affix									building	
ing	Affix									flying	
ing	Affix									hosting	
ing	Affix									meaning	
ing	Affix									painting	
ing	Affix									thinking	
ing	Affix									warning	

Merge Show Filtered

Merging Morphemes

To merge morphemes, select all the morphemes/allomorphs to be merged and click the **Merge** button located beneath the Morpheme Explorer. This automatically saves your work.

Merged Morpheme in the Morpheme Explorer Window

Morpheme Explorer											
Morpheme	Type	Part of Speech	Tense	Polarity	Person	Class	Number	Gender	Other Attributes	Examples	
ing	Affix									boiling, building, flying, hosting, m.	

To set morpheme attributes, select all the desired morphemes and set their attributes in the [Word/Morpheme Information](#) frame.

Show Filtered

In order to view your morpheme assignments, you can highlight the relevant morpheme(s) and click on the **Show Filtered** button. This filters out all words in the Word Collection table that do not contain the selected morpheme(s) that you have highlighted in the Morpheme Explorer window. This allows you to easily review and modify any morpheme assignments you have previously made in your collection.

Clearing Cuts

To clear cuts (i.e., remove a previously assigned morpheme status), select the morpheme in the Morpheme Explorer and click **Show Filtered**. Once the group of words appears in the Word Collection table, highlight the applicable morphemes (which are designated as cuts by appearing as either gold or silver for affixes and bronze or copper for roots) and click the **C** button to the right of the table. This will clear your cuts and return the morphemes back to white (which re-designates them as undefined substrings).

Word/Morpheme Information

Once you have merged related morphemes together, click on the morpheme to which you would like to assign specific morpheme information. To the left of the Morpheme Explorer, the Morpheme Attributes window will become active so that you can add relevant attributes.

Morpheme Attributes

In each category below, you have the option of selecting *Other*. This brings up a text field where you can write in your own information or value for a particular category. Information you add about a morpheme will automatically be saved.

Morpheme Attribute Window

The image shows a software window titled "Word/Morpheme Information". It has two tabs: "General" and "Morpheme Attributes". The "Morpheme Attributes" tab is selected. Inside this tab, there are eight dropdown menus arranged in a 4x2 grid. The first dropdown, "Type", is set to "Affix". The other dropdowns are "Part of Speech", "Tense", "Polarity", "Person", "Class", "Number", and "Gender", all of which are currently empty. Below this grid is a text input field labeled "Other Attributes".

Category	Description	Options
Type	Sets the type of all morphemes selected in the Morpheme Explorer. A morpheme's type is automatically set to either <i>Root</i> or <i>Affix</i> , depending on the command used to create the morpheme. Here, you can be more specific.	Root, Affix, Suffix, Prefix, Infix, Other
Part of speech	Sets the part of speech of all morphemes selected in the Morpheme Explorer.	Adjective, Noun, Proper noun, Verb, Other
Tense	Sets the tense of all morphemes selected in the Morpheme Explorer.	Present, Past simple, Past imperfect, Subjunctive, Other
Polarity	Sets the polarity of all morphemes selected in the Morpheme Explorer.	Affirmative, Negative
Person	Sets the person attribute of all morphemes selected in the Morpheme Explorer.	1 st , 2 nd , 3 rd , Other
Class	Sets the class of all morphemes selected in the Morpheme Explorer.	Class numbers through 20 are given, Other
Number	Sets the number value of all morphemes selected in the Morpheme Explorer.	Singular, Dual, Plural, Other
Gender	Sets the gender of all morphemes selected in the Morpheme Explorer	Feminine, Masculine, Neutral, Other

<i>Category</i>	<i>Description</i>	<i>Options</i>
Other Attributes	<p>Provides a way to add user-defined, customized morpheme attributes to each morpheme selected in the Morpheme Explorer.</p> <p>While data can be entered in whatever way best suits your needs, we suggest delineating each attribute and using a category-value pair. For example, one way to add 3 categories (A, B, C) with respective values (va, vb, vc) is to pair the category and value with an equal sign, delineated with a semi-colon or other punctuation (e.g., A=va; B=vb; C=vc). This kind of delineation makes it both human and computer readable. Since Alchemist documents are saved in XML format, a semi-colon delineator is safe and readable by any program accessing the saved file.</p>	N/A (text-entry box)

General Word/Morpheme Information

In this window, you can assign information to individual words or morphemes. For example, if you assign status to morphemes for a group of words but are less certain about some words than others in the group, you can change the default *Certain* score to reflect your (un) certainty and then add any notes or remarks about it.

Scoring is done on a per morpheme basis. The scoring options are *Certain*, *Not scored*, *Somewhat certain*, or *Uncertain*, which you assign to reflect the certainty that a cut morpheme is correct. The score for all words (and partitions) is defaulted to *Not scored* before any cuts are made and is automatically changed to *Certain* for all new cuts. Using this field will overwrite the score of every selected morpheme, even if only one letter of a morpheme is selected.

The Notes/Remarks feature allows you to add notes/remarks and keep track of your comments about morpheme assignments. There are no particular length restrictions or requirements for this field. The notes/remarks are per morpheme, as opposed to per word, but can be applied to an entire word by selecting all morphemes of the word before writing the notes. Using this field will overwrite the notes/remarks of every selected morpheme, even if only one letter of a morpheme is selected.

To save the data you enter into the Notes/Remarks section for the selected word(s) or morpheme(s), click the **Commit** button. To view scoring or notes, click any part of the word or morpheme in the Word Collection table.

Saving Data

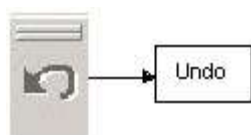
To save the collection to a pre-chosen XML (*.xml) file, go to the File menu and click **Save** (shortcut Ctrl+S) or click on the Save icon (located under the Morphology menu). If the file has not been previously saved or you want to save it to a new XML (*.xml) file, go to the File menu and click **Save As** and follow the prompts to name and save it.

Guide to Navigation Tabs and Menus

Navigation Tabs

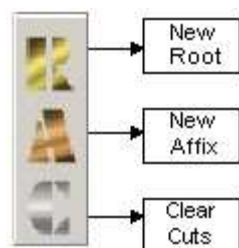
Undo

The button appearing at the top right corner allow you to undo a previous deletion (before you have clicked **Commit Deletions**).



RAC

The buttons appearing under the Undo button allow you to assign a new root (shortcut key Ctrl+R) or new affix (shortcut key Ctrl+A) to the words in your collection, as well as clear any cuts you have made but do not want to assign (shortcut key Ctrl+C). Root assignments will be gold and silver, affix assignments will be bronze and copper, and undefined substrings will be white.



Navigation Menus

File

<i>Option</i>	<i>Description</i>	<i>Shortcut</i>
New	Allows you to open a new text (*.txt) file to feed into your collection. When you open a new file, you will be prompted with the question, "How many distinct words would you like in the gold standard collection?" The default number is 500.	Ctrl + N
Open	Allows you to open a working XML (*.xml) file.	Ctrl + O
Save	Allows you to save all words from the collection to a pre-chosen XML (*.xml) file. The XML file can be reopened and edited in Alchemist.	Ctrl + S

<i>Option</i>	<i>Description</i>	<i>Shortcut</i>
Save as	Allows you to save the words in your collection to a new XML (*.xml) file. The XML file can be reopened and edited in Alchemist.	
Read More Words	Allows you to add more words to an open file that you are viewing in the Word Collection table.	
Close	Allows you to close the currently open file.	
Export All	Exports a text (*.txt) file that includes all the words in your Word Collection table.	
Export Filtered	Exports a text (*.txt) file that includes only the filtered words displayed in the Word Collection table.	
Export Wizard	Brings up a wizard dialog that allows users to format the exported data and export to a text file.	
Preferences	Allows you to specify how many words you would like to view per page. Words in your collection are displayed in pages, which are sets of words of up to 200 in number. Page through the set of words using the controls in the lower right (e.g., the Previous Page and Next Page tabs, and the text-entry box that allows you to choose a particular page).	
Exit	Allows you to exit the Alchemist program.	

Edit

<i>Option</i>	<i>Description</i>	<i>Shortcut</i>
Undo Deletions	Allows you to undo your most recent action.	Ctrl + Z
Commit Deletions	Allows you to refresh your window in order to reflect any deletions you have made.	F5
Delete	Allows you to delete data from selected cells in the Word Collection table. Highlight the entire row or rows you want to delete by clicking on the number(s) associated with the word(s). You can also delete rows using the Delete tab located on keyboard. If all cells of a word are selected, the entire word will be deleted from the collection. Otherwise, only the part of the word that is selected will be removed.	
Scrub	Allows you to clean up text files before analyzing words in the Word Collection table. There are simple and advanced scrubbing options in the Word Scrubbing window.	Ctrl + B
Font	Allows you to set your font preferences (e.g., font name, font style, size, effects, and script). Keep in mind that roots will continue to appear in bold and affixes will appear in italics, regardless of the font preferences you select.	Ctrl + F

<i>Option</i>	<i>Description</i>	<i>Shortcut</i>
Left to right	Toggles the directionality of the text (i.e., individual characters in the columns). Text in the words column will automatically be written with a language's correct directionality (e.g., left-to-right for Romance languages, right-to-left for Semitic languages.)	
Document Information	Opens a dialog where the creator name, affiliation, corpus language and general remarks can be edited. This allows the author to provide meta-data to the document so other people using the files can give proper credit where due.	

Morphology

<i>Option</i>	<i>Description</i>	<i>Shortcut</i>
New Root	Allows you to assign a new root to your collection.	Ctrl + R
New Affix	Allows you to assign a new affix to your collection.	Ctrl + A
Clear Cuts	Allows you to clear (or remove) any assignments you have previously made to a word or words but do not want to include in your collection analysis.	Ctrl + C
Duplicate Word	Allows you to duplicate a word. This is useful if you want to record a word in more than one way. For example, the word 'polish' in English can be analyzed as both 'pol+ish' as in 'being from Poland' or as the verb 'polish.' Creating a duplicate word allows you to account for both entries.	Ctrl + D
Merge Allomorphs	Allows you to group all selected morphemes and/or allomorphs into one entry in the Morpheme Explorer.	

Help (shortcut F1):



Once you click on this button, you can click on any feature in Alchemist to learn more about what it does and how to use it.