1   Graphical User Interface

In this section, we describe the functionalities of the graphical user interface provided in our toolkit and present some examples. The interface consists of three panels, which correspond to the three main functionalities: sentence generation, sentence enumeration and sentence analysis. Each of these panels is implemented independently as a `Java` class, so that each of them can be used independently. A general view of the interface is presented in Figure 1.

1.1   Editor and Console

Each panel contains a grammar editor which enables the user to edit, load or save a grammar and to check whether a grammar is well-defined, and an output console which outputs the outcomes and extra information like sentence length, running time and so on.

The grammar rules are written as the following form: $<X>->\alpha_1|\ldots|\alpha_n$, where nonterminals are bracketed by angle brackets to be distinguished from terminals. Once the grammar rules are input, the editor will check whether the grammar is well-defined. Figure 1 shows an example of editing the sample grammar. By default, the start symbol corresponds to the left-hand side of the first rule. The user can also indicate the start symbol through the "Start" dropdown box.

1.2   Sentence Generation Panel

The sentence generation panel allows the user to generate a sentence set that satisfies rule coverage or context-dependent rule coverage. A length control mechanism can as well be configured on in the sentence generation process by marking the "Length Control" check box. After generation, the number of generated sentences, the information of sentences' lengths, and the total time for generating sentences are also counted. For instance, we generate a sentence set that satisfies rule coverage from the sample grammar with the length control. The result is shown in Figure 2.

1.3   Sentence Enumeration Panel

The sentence enumeration panel consists of three tabs. Figure 4 shows a general view of the panel.

The first tab enables the user to query the volume data among a bounded hierarchical range by giving the "From" hierarchy index and the "To" hierarchy index. It also allows the user to choose which kinds of volume data to query by marking the corresponding check boxes. For example, Figure 3 shows the result about querying the hierarchy and heap volume data of the sample grammar from the 5-th hierarchy to the 7-th hierarchy.

The second tab allows the generation of the previous or next sentence for a given sentence and the conversion between sentences and natural numbers.
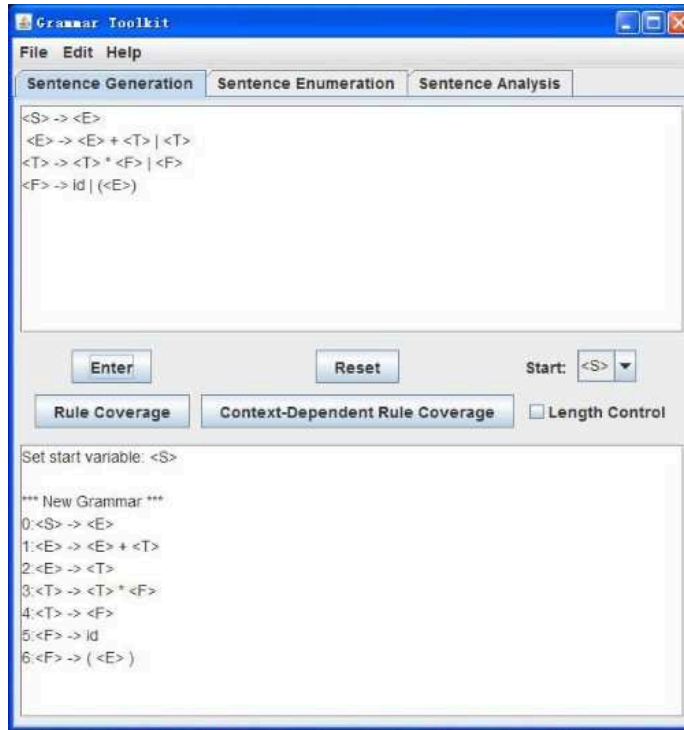
Figure 1: Editing of the sample grammar

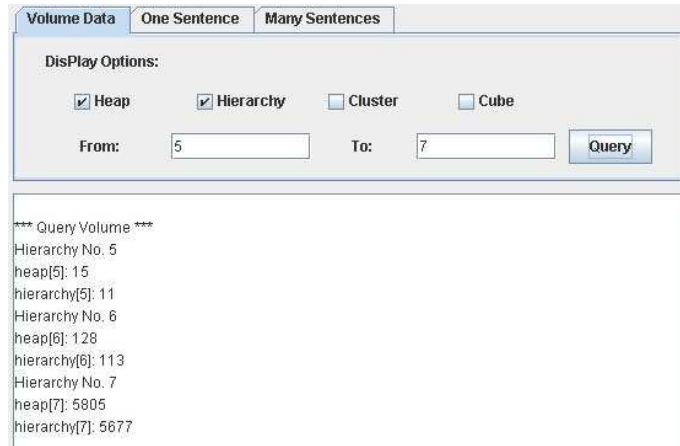Figure 2: Sentence generation with rule coverage and length control

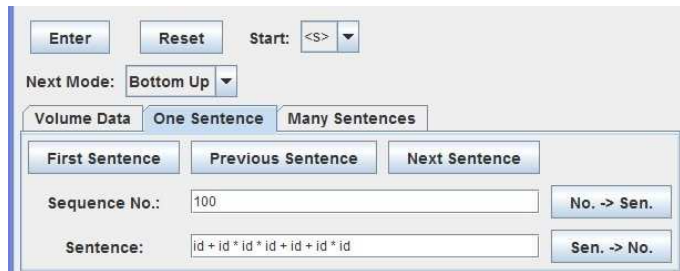Figure 3: The hierarchy and heap volume data from the 5-th to 7-th hierarchy

Figure 4: Computing the 100th sentence

There are two modes, bottom-up and top-down, for computing the next or previous sentence of a given sentence. Through the "Next Mode" dropdown box, the user can choose either he likes. According to hierarchical lexicographic order, there exists a one-to-one correspondence between natural numbers and sentences. Therefore, given a sentence, we can compute its sequence number and vice versa. An example is to compute the 100-th sentence of the sample grammar, showed by Figure 4.

The third tab allows the user to enumerate a bounded range of sentences by giving the sequence numbers of the first and last sentences. The range can be restricted in a hierarchy, cluster or cube, by giving the corresponding indices in which case only the sentences in the intersection of the range and the restriction are enumerated. The hierarchies, clusters and cubes are used for expert users. For normal users, we will consider how to give the restriction by giving some rules or structures in the future. After enumeration, the total time for enumerating sentences, the actual number of enumerated sentences and the average time for enumerating a sentence are also counted. For instance, we enumerate the sentences in the range [0,100] in the 5-th hierarchy for the sample
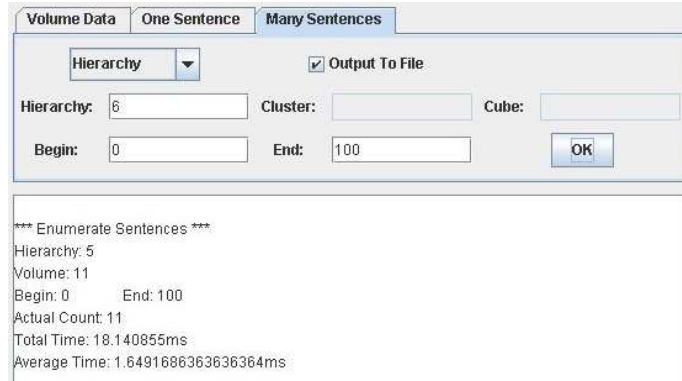
3

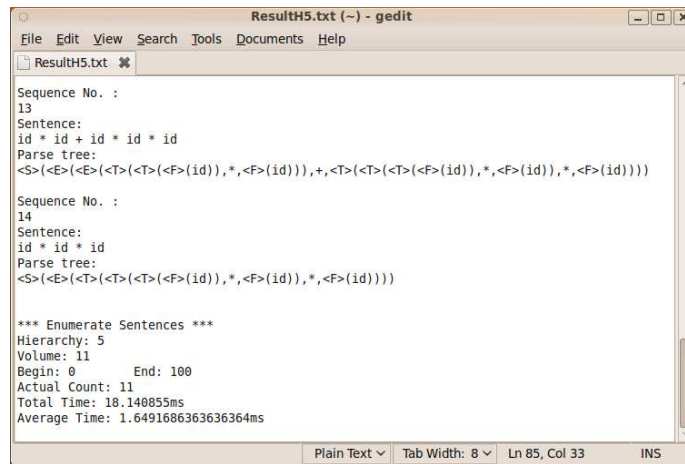Figure 5: Enumerating the sentences in [0,100] in the 5-th hierarchy

Figure 6: The output file containing the objects in [0,100] in the 5-th hierarchy

grammar. Since the 5-th hierarchy volume data is 11 (< 100), only 11 sentences are enumerated. The result is given in Figure 5.

When enumerating a bounded range of sentences, the console does not output all the sentences. Additionally, the toolkit allows the user to export the enumerated sentences into a file if he sets the "File" checkbox on. The output file contains the corresponding sequence number and parse tree for each enumerated sentence, as well as the statistic information mentioned above. Figure 6 presents the output file of the above example (See Figure 5).

1.4 Sentence Analysis Panel

Through the sentence analysis panel, the user can parse a sentence or analyze the coverage of a sentence set, including rule coverage and context-dependent

4

Figure 7: Analyzing the coverage of a sentence set generated from the sample grammar

rule coverage. Since we have generated a sentence set from the sample grammar with rule coverage and length control (see Figure 2), we parse this sentence set as an instance. Figure 7 shows the result of the coverage analysis for the sentence set.