Frank Drewes
drewes@cs.umu.se
+46 90 786 9790

UMEÅ UNIVERSITY

Overall Problem Setting and Grading Criteria

The assignment part of the course consists of individual work in which algorithms are developed, implemented, and their efficiency is analyzed. The major output, and the basis for the grade achieved, is a scientific report that describes the work and its results.

1 The Task

The task revolves around the membership problem for context-free grammars, which can be solved efficiently by the Cocke-Younger-Kasami (CYK) algorithm, a typical example of dynamic programming. Students are advised to read the corresponding sections of a good textbook on context-free grammars and the CYK algorithm; *Introduction to Automata Theory, Languages, and Computation* by Hopcroft and Ullman (1979) is the recommended reference, but there are numerous others.

The membership problem asks, for a given context-free grammar and an input string, whether this input string is an element of the language generated by the grammar. The original CYK algorithm is a dynamic programming algorithm that solves the problem in a bottom-up fashion. Here, we implement and analyse this original version, but also the naive version that does not make use of dynamic programming, and the top-down version using memoization. We also modify the algorithm to solve easier or more general problems, investigating whether one can gain efficiency in the first case and avoid to loose too much in the second case.

2 Four Steps

The work is broken down in four steps. These steps are described in more detail in four separate assignments published on the course homepage. The description below is only intended to provide a rough orientation. Handing in the results of steps 1–3 in order to receive feedback is voluntary but *highly* recommended. The rules for receiving feedback on step i are the following: the previous steps must have been handed in and received the grade *pass* (G), and step i must have been handed in by the specified deadline. Last but not least, the submission must be reasonably complete and of sufficient quality. If these conditions are satisfied, feedback will be provided (which should be used to make improvements) and the submission will be graded *pass*. Note that, even if the grade is *fail* (U), the final report should be handed in and will be graded just like any other.

2.1 Implementation An implementation of the basic algorithms and the infrastructure required for the empirical tests is made. The implementation makes use of Java and has to display reasonable behaviour in terms of correctness and efficiency.

2.2 Initial Description and Analysis The implementation is used in order to perform an initial evaluation of the efficiency of the algorithms, exploring several cases to see whether different inputs favour different variants of the algorithm, and trying to analyze why. A first version of the scientific report is written which introduces the notions, the problem setting, the results achieved, and the analysis. Draft reports handed in by the deadline will be commented on by the course staff, provided that they are sufficiently complete and it is visible that a reasonable effort has been spent on them.

2.3 Special Cases and Generalizations Special cases as well as generalizations are explored. For this, the implementations are modified accordingly. The special cases concern

(a) the so-called linear grammars and
(b) generalizations that count how many errors an input string contains.

Students solve (a) for a maximum grade 3, (b) for a maximum grade 4, and both for a maximum grade 5. (Here, "maximum" means that, naturally, the grade also depends on the quality of the solution; see below.) Again, the efficiency is analyzed, and the report is extended accordingly. Again, draft reports handed in by the deadline will be commented on by the course staff under the same conditions as above.

Frank Drewes
drewes@cs.umu.se
+46 90 786 9790

UMEÅ UNIVERSITY

2.4 Finalizing the Report The last step is to finalize the report. This report, together with the implementation, is what will finally be graded.

3 The Implementation and Report

The final report is the major output and the basis for the grade. The implementation, in turn, is the basis for the findings in the report. Make sure to implement things in a simple yet efficient manner and check that the implementation works as it should, because not doing so is likely to result in a considerable waste of time. (This *has* happened frequently in the past, in several cases to an almost ridiculous extent!)

Here is a very important hint not just for this assignment but for empirical evaluation of algorithms in general:

> *Be skeptical and avoid blind trust in the implementation. Always ask yourself whether the results make sense and can be explained by reasoning. They can both be too good and too bad. If in doubt, investigate the matter until an error in the implementation or a good explanation has been found. Be especially critical when comparing empirical results (based on implementations) with theoretical expectations. If there are (apparent or real) discrepancies, find out why.*

As mentioned, the major output is the report. It is to be written in English, in the style of a self-contained scientific text rather than a colloquial exposition. Thus, it should have an abstract, a reasonable introduction, and a good logical structure, making it easy to read and understand the purpose of each part within the whole.

Make sure that the text makes sense even for readers not involved in the course. The best way to achieve this is to pretend that the task is not an assignment which is part of a course. Rather, it is to investigate a scientific question (which is actually the case!) and the goal is to convey the findings to the reader. In particular, do not make implicit references to the specification of the assignment based on the assumption that the reader knows the context of the text.

Make sure to argue convincingly when, e.g., discussing advantages and disadvantages of algorithmic techniques or analyze the efficiency of algorithms. Would you be convinced by the arguments yourself if someone else would try to convince you? A necessary condition for this is to explicitly mention all major assumptions on which the arguments or decisions are based. A very common (and hard to avoid) mistake is to confuse readers with mind readers.

References are an important part of scientific work, serving several purposes. They inform the reader about the origin of important terms, notions, and results. They can be used for disambiguation, if terms are used in different ways in the literature. They can direct the reader to additional information and related work not explicitly discussed. Finally, the by far most important use of references is that they lend support to the reasonings and arguments. If the reasoning aligns with other people's work it becomes more credible. However, do not make the mistake to adopt uncritically whatever may be found in the literature. The citation of a reference as support for some argument also means that one agrees with whatever is cited (unless a reference is cited explicitly in order to direct attention to a discrepancy or mistake).

Avoid relying on Wikipedia and similarly unstable resources. They are usually a good way to come up with ideas and as a start for a search for reliable material, but are too volatile and uncertain to rely upon as major sources. Instead, make sure to identify reliable sources and refer to them.

4 Grading and Grading Criteria

The grade will be based on an evaluation of the scientific quality of the report, according to the following criteria:

Frank Drewes
drewes@cs.umu.se
+46 90 786 9790

UMEÅ UNIVERSITY

Structure, language, and writing style Is the report well structured, so that the division into sections and paragraphs makes sense and guides the reader through the report in an appropriate way? Is the language correct and the style scientific in nature? Are examples, illustrations, and explanations used in a way that supports understanding?

Argumentation and analysis Does the contents make sense, does it ask interesting questions, and are the answers argued for in a convincing manner, without major parts being missing? Is speculation clearly separated from hard facts and definite results?

Correctness Are the definitions and arguments as well as the described algorithms correct, are the experimental results and their analysis reasonable, and is the implementation itself sufficiently clean to be trustworthy?

Completeness Does the report contain all the required parts? The highest grade can only be achieved if both (a) and (b) in Step 3 above are included. If (a) resp. (b) is missing, the highest achievable grade is 4 resp. 3.

For a passing grade, the quality with respect to each of the four aspects above must be decent. Regarding the completeness aspect, this means that it is insubstantial for the distinction between U on the one hand and 3, 4, 5 on the other hand whether (a) and (b) are both included or only one of them is. What is important is the quality of the parts that *are* included. Which of the two analyses are included discriminates exclusively between the passing grades. If in doubt, one should thus opt for quality rather than greater coverage.[1]

If the report receives the grade fail (U), it can be improved and resubmitted by a date to be announced. If that is still graded U (or no solution at all has been handed in), there is a third attempt that can be made in spring, when the 3rd course period ends.

[1] An aspect in favor of (b) is that it is the more challenging and interesting task.