SG2 User Manual

(CS 4500, Group 5)

Class: CS 4500
Project: SG2 User Manual
Group: Group 5
Team Members:

- Caleb Jacobs
- Cynthia Brown
- DeJa Thompson
- Dheaven Hill
- Gabe Jackson

Table of Contents

1. Introduction

SG2 is a text-processing program designed for CS 4500. It reads multiple .txt files, pulls out all their words, counts how often each word appears, and helps you search for specific words. Then it generates two reports:

- **Concordance.txt** – lists every word and exactly where it appears.
- **ExtraLists.txt** – lists word frequencies, shared words, and unique words.

SG2 builds on SG1 but adds several new features, which are explained in Section 5.

---

2. What SG2 Does

SG2 handles up to **10 text files at once**. For each file, it:

- Reads every word
- Fixes words broken by hyphens across lines (like "eco-" + "friendly")
- Counts total and distinct words
- Lets you search for words across all loaded files
- Creates two reports automatically when done

SG2 is meant to help students understand text parsing, file input, and frequency mapping.

---

3. Getting Started

Before using SG2:
- Make sure your .txt files are in the same folder as the program.
- Only files ending in .txt (any casing, like .txt, .txt) are allowed.
- Word searches must follow the program's rules: letters only, with optional **internal hyphens**.

How to Run SG2
1. Start the program.
2. When asked, enter a .txt filename (ex: sample1.txt).
3. Choose whether to load more files.
4. Begin searching for words.

---

4. How to Use SG2

Loading Files

- The program will ask for a .txt filename.
- After each file, it will ask: **"Add another file?"**
- If you enter the same file twice, SG2 blocks duplicates.

Searching for Words

Words must:

- Use letters A–Z
- Have **internal hyphens only** (ex: first-base)
- Not start or end with a hyphen
- Not contain double hyphens (–)

After each search, SG2 asks if you'd like to search again.

Reports Generated

1. Concordance.txt
Lists every word and its exact positions in the format:
(fileNumber.lineNumber.wordNumber)

2. ExtraLists.txt
Includes:

- Top 10 most frequent words
- Words appearing in all loaded files
- Words unique to a single file

5. SG2 New Features (Compared to SG1)

This section highlights improvements you now get in SG2.

Multiple File Support

You can now load **up to ten files**, instead of just one.

Duplicate File Protection

SG2 checks if you accidentally enter the same filename twice.

Cross-File Searching

You can search for a word across **all loaded files at once**, not just one.

More Detailed Reports

SG2 adds:

- Shared words across all files
- Words that appear in only one file
- Top 10 frequency list

Improved Handling of Hyphenated Lines

If a word is split like this:

eco-

friendly

SG2 recognizes it as **"eco-friendly"**.

6. Troubleshooting

Problem: "File Not Found"

- Make sure your .txt file is in the same folder as the Python program.

Problem: Word is "Illegal"

Check that your word:

- Contains only letters and INTERNAL hyphens
- Does not start/end with a hyphen
- Has no spaces or symbols

Problem: Program Stops Early

Make sure to answer **Yes/No** exactly as allowed:

- Yes, yes, y, Y
- No, no, n, N

7. Revision History

This manual is based on SG2 and A revised version of SG1. Here are the dates of the revisions made:

10/14/2025 - Joel generated SG0.5

10/14/2025 - SG1 is ready for customer review.

11/09/2025- Initial SG2 Revision

11/13/2025 SG2 revision is ready for customer review.