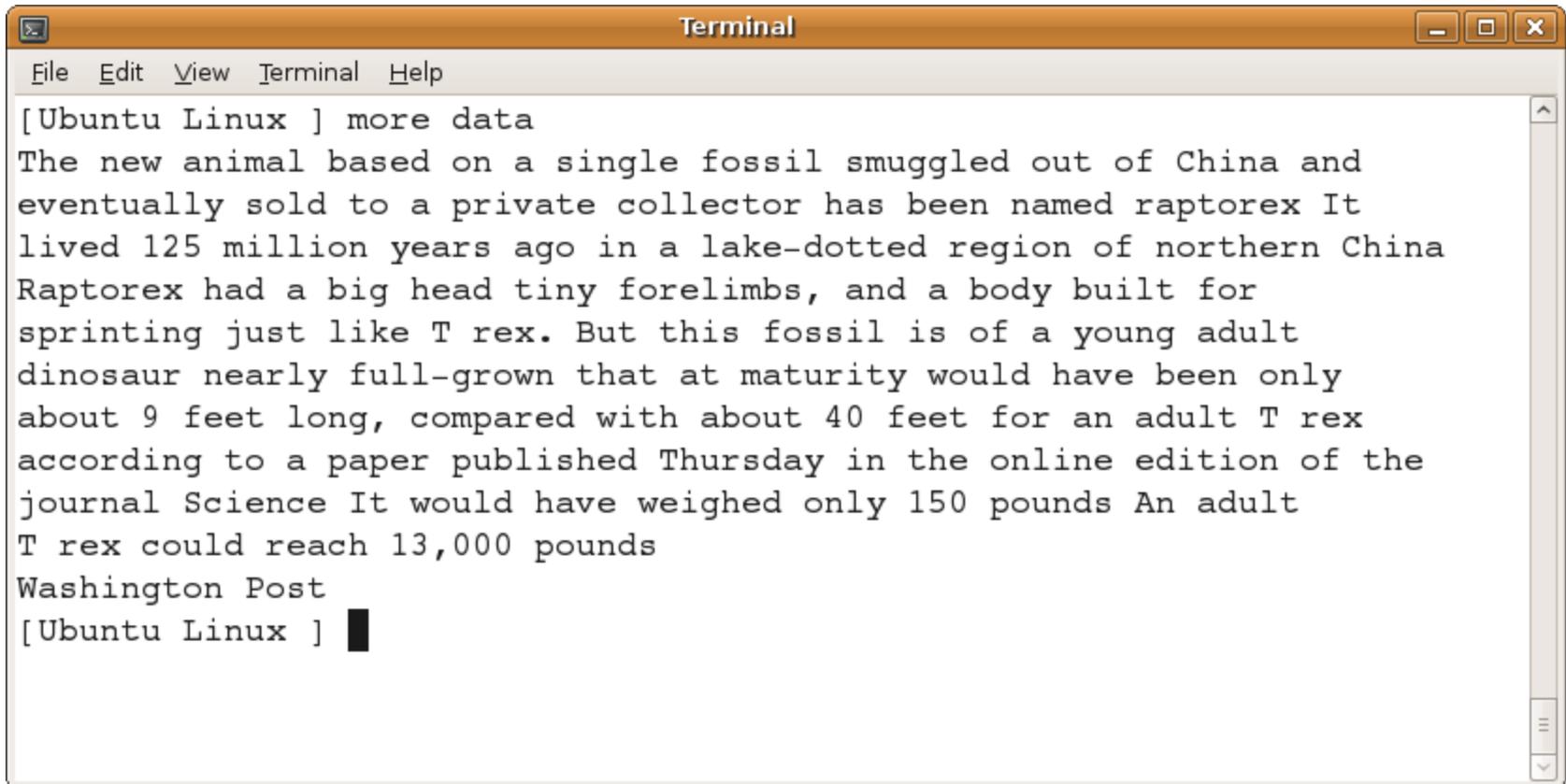


Sorting a File Line by Line Using the n-th Words

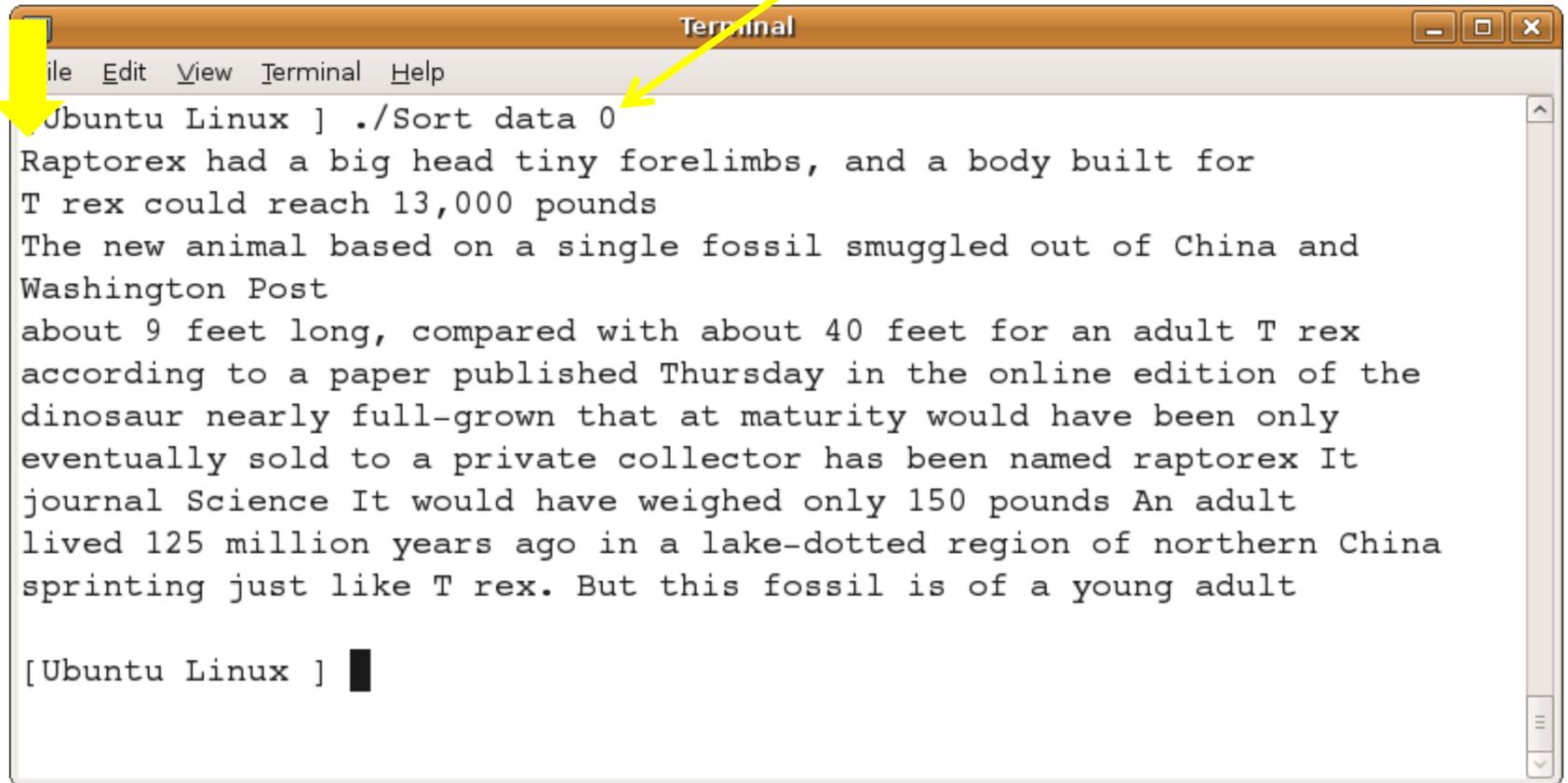
Yung-Hsiang Lu

Input

A terminal window titled "Terminal" with a menu bar containing "File", "Edit", "View", "Terminal", and "Help". The terminal content shows a prompt "[Ubuntu Linux]" followed by the input "more data". The output is a paragraph of text about a dinosaur fossil named Raptorex, followed by the source "Washington Post". The terminal ends with a prompt "[Ubuntu Linux]" and a black cursor block.

```
[Ubuntu Linux ] more data
The new animal based on a single fossil smuggled out of China and
eventually sold to a private collector has been named raptorex It
lived 125 million years ago in a lake-dotted region of northern China
Raptorex had a big head tiny forelimbs, and a body built for
sprinting just like T rex. But this fossil is of a young adult
dinosaur nearly full-grown that at maturity would have been only
about 9 feet long, compared with about 40 feet for an adult T rex
according to a paper published Thursday in the online edition of the
journal Science It would have weighed only 150 pounds An adult
T rex could reach 13,000 pounds
Washington Post
[Ubuntu Linux ] █
```

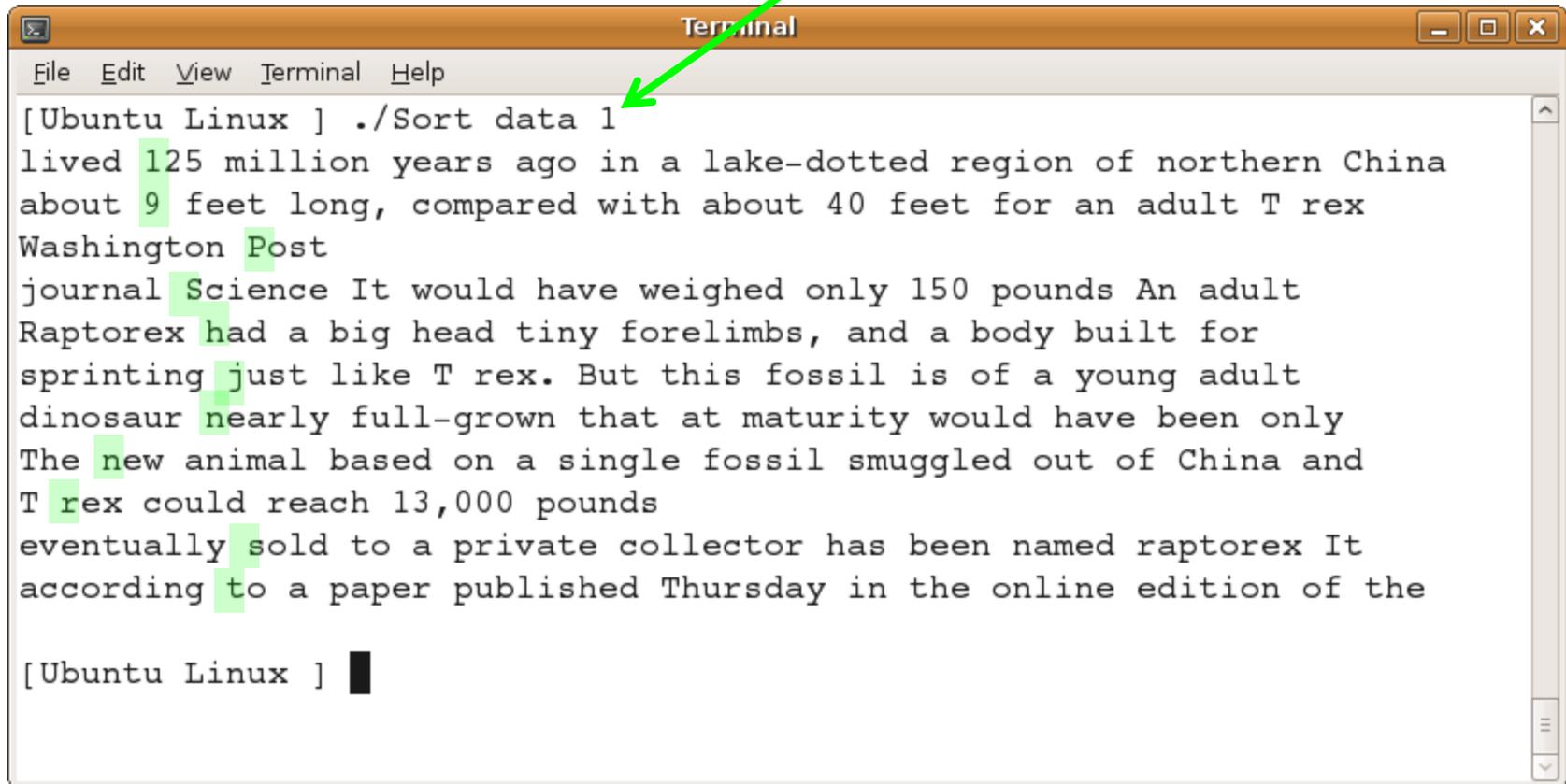
Sorting by the First Word

A terminal window titled "Terminal" with a menu bar containing "File", "Edit", "View", "Terminal", and "Help". The prompt is "[Ubuntu Linux]". A yellow arrow points from the word "First" in the title to the command line. The command is `./Sort data 0`. The output is a block of text about a dinosaur fossil. A second yellow arrow points from the word "Sort" in the command to the first word of the output, "Raptorex".

```
[Ubuntu Linux ] ./Sort data 0
Raptorex had a big head tiny forelimbs, and a body built for
T rex could reach 13,000 pounds
The new animal based on a single fossil smuggled out of China and
Washington Post
about 9 feet long, compared with about 40 feet for an adult T rex
according to a paper published Thursday in the online edition of the
dinosaur nearly full-grown that at maturity would have been only
eventually sold to a private collector has been named raptorex It
journal Science It would have weighed only 150 pounds An adult
lived 125 million years ago in a lake-dotted region of northern China
sprinting just like T rex. But this fossil is of a young adult

[Ubuntu Linux ] █
```

Sorting by the Second Word

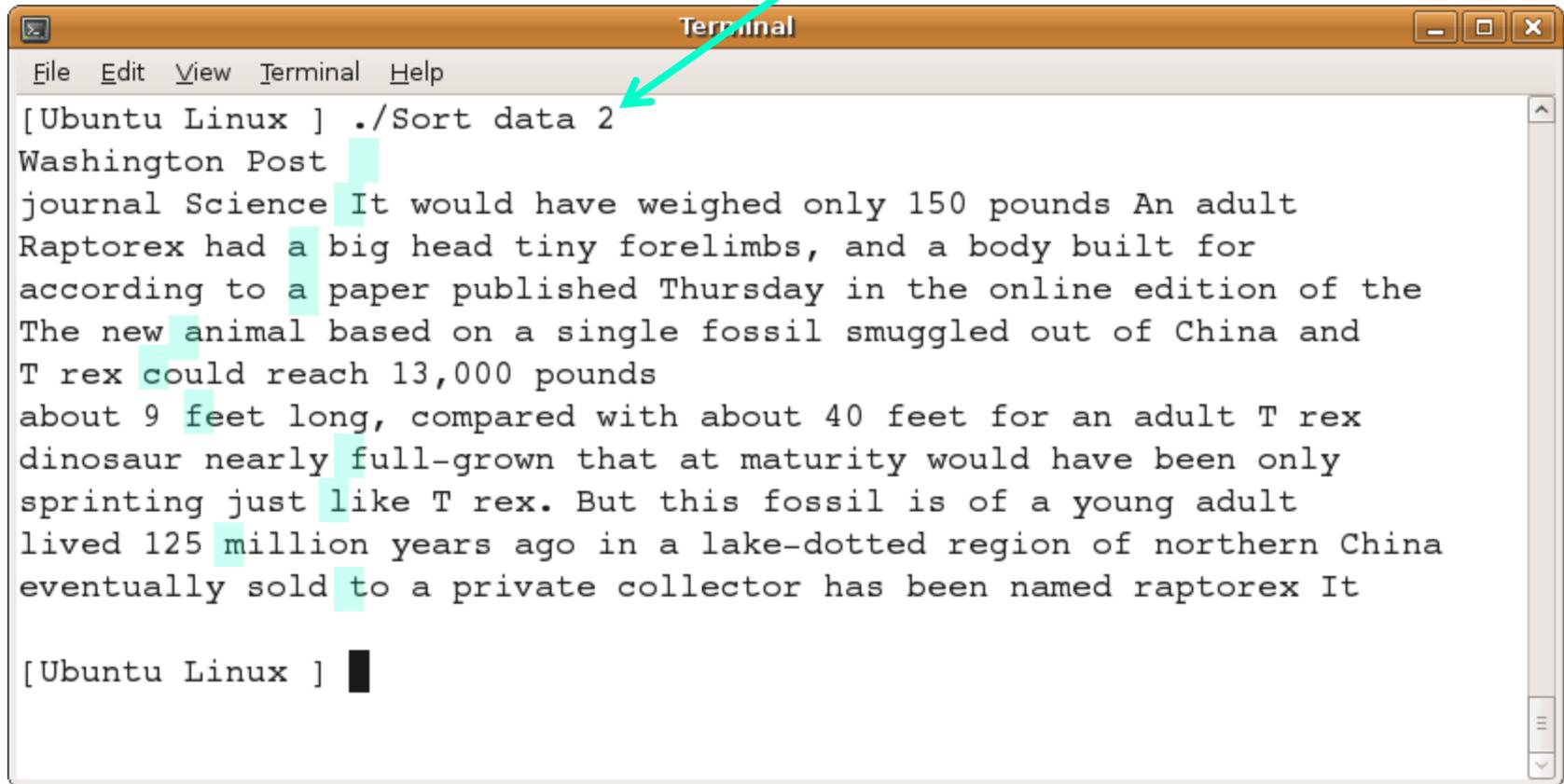


A terminal window titled "Terminal" with a menu bar containing "File", "Edit", "View", "Terminal", and "Help". The terminal shows the command `./Sort data 1` being executed. The output is a list of text lines, some of which are highlighted in green. A green arrow points from the word "Second" in the title above to the number "1" in the command. The output lines are: "lived 125 million years ago in a lake-dotted region of northern China", "about 9 feet long, compared with about 40 feet for an adult T rex", "Washington Post", "journal Science It would have weighed only 150 pounds An adult", "Raptorex had a big head tiny forelimbs, and a body built for", "sprinting just like T rex. But this fossil is of a young adult", "dinosaur nearly full-grown that at maturity would have been only", "The new animal based on a single fossil smuggled out of China and", "T rex could reach 13,000 pounds", "eventually sold to a private collector has been named raptorex It", "according to a paper published Thursday in the online edition of the". The terminal ends with a prompt `[Ubuntu Linux]` and a black cursor block.

```
[Ubuntu Linux ] ./Sort data 1
lived 125 million years ago in a lake-dotted region of northern China
about 9 feet long, compared with about 40 feet for an adult T rex
Washington Post
journal Science It would have weighed only 150 pounds An adult
Raptorex had a big head tiny forelimbs, and a body built for
sprinting just like T rex. But this fossil is of a young adult
dinosaur nearly full-grown that at maturity would have been only
The new animal based on a single fossil smuggled out of China and
T rex could reach 13,000 pounds
eventually sold to a private collector has been named raptorex It
according to a paper published Thursday in the online edition of the

[Ubuntu Linux ] █
```

Sorting by the Third Word

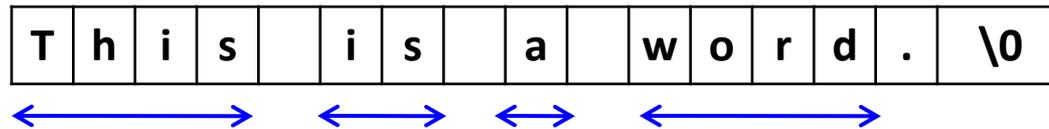


A terminal window titled "Terminal" with a menu bar containing "File", "Edit", "View", "Terminal", and "Help". The terminal shows the command `./Sort data 2` being executed. The output is a text file with the following content: "Washington Post", "journal Science", "It would have weighed only 150 pounds", "An adult Raptorex had a big head tiny forelimbs, and a body built for according to a paper published Thursday in the online edition of the The new animal based on a single fossil smuggled out of China and T rex could reach 13,000 pounds about 9 feet long, compared with about 40 feet for an adult T rex dinosaur nearly full-grown that at maturity would have been only sprinting just like T rex. But this fossil is of a young adult lived 125 million years ago in a lake-dotted region of northern China eventually sold to a private collector has been named raptorex It". The terminal prompt `[Ubuntu Linux]` is visible at the end of the output.

```
[Ubuntu Linux ] ./Sort data 2
Washington Post
journal Science
It would have weighed only 150 pounds
An adult Raptorex had a big head tiny forelimbs, and a body built for
according to a paper published Thursday in the online edition of the
The new animal based on a single fossil smuggled out of China and
T rex could reach 13,000 pounds
about 9 feet long, compared with about 40 feet for an adult T rex
dinosaur nearly full-grown that at maturity would have been only
sprinting just like T rex. But this fossil is of a young adult
lived 125 million years ago in a lake-dotted region of northern China
eventually sold to a private collector has been named raptorex It

[Ubuntu Linux ] █
```

For simplicity, we assume two adjacent words are separated by one and only one space.



```
C/C++ - sort.c - Eclipse SDK
File Edit Navigate Project Window Help
sort.c
char * findNthword(char * line, int n)
{
    /* return the pointer to the n-th word */
    /* If the line does not have n words, return NULL. */
    /* consider only one space separating two adjacent words */
    int wordCount = 0;
    char * wordPtr = line;
    while ((wordPtr != NULL) && (wordCount < n))
    {
        wordPtr = strchr(wordPtr, ' ');
        if (wordPtr != NULL)
        { wordPtr++; }
        wordCount++;
    }
    return wordPtr;
}
```

Writable Smart Insert

Find the Third Word (n is 2)

wordPtr = line ↓

| | | | | | | | | | | | | | | | |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|
| T | h | i | s | | i | s | | a | | w | o | r | d | . | \0 |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|

`wordPtr = strchr(wordPtr, ' ');`

line ↓ **wordPtr** ↓

| | | | | | | | | | | | | | | | |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|
| T | h | i | s | | i | s | | a | | w | o | r | d | . | \0 |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|

`wordPtr ++;`

line ↓ **wordPtr** ↓

| | | | | | | | | | | | | | | | |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|
| T | h | l | s | | i | s | | a | | w | o | r | d | . | \0 |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|

`wordPtr = strchr(wordPtr, ' ');`

line ↓ **wordPtr** ↓

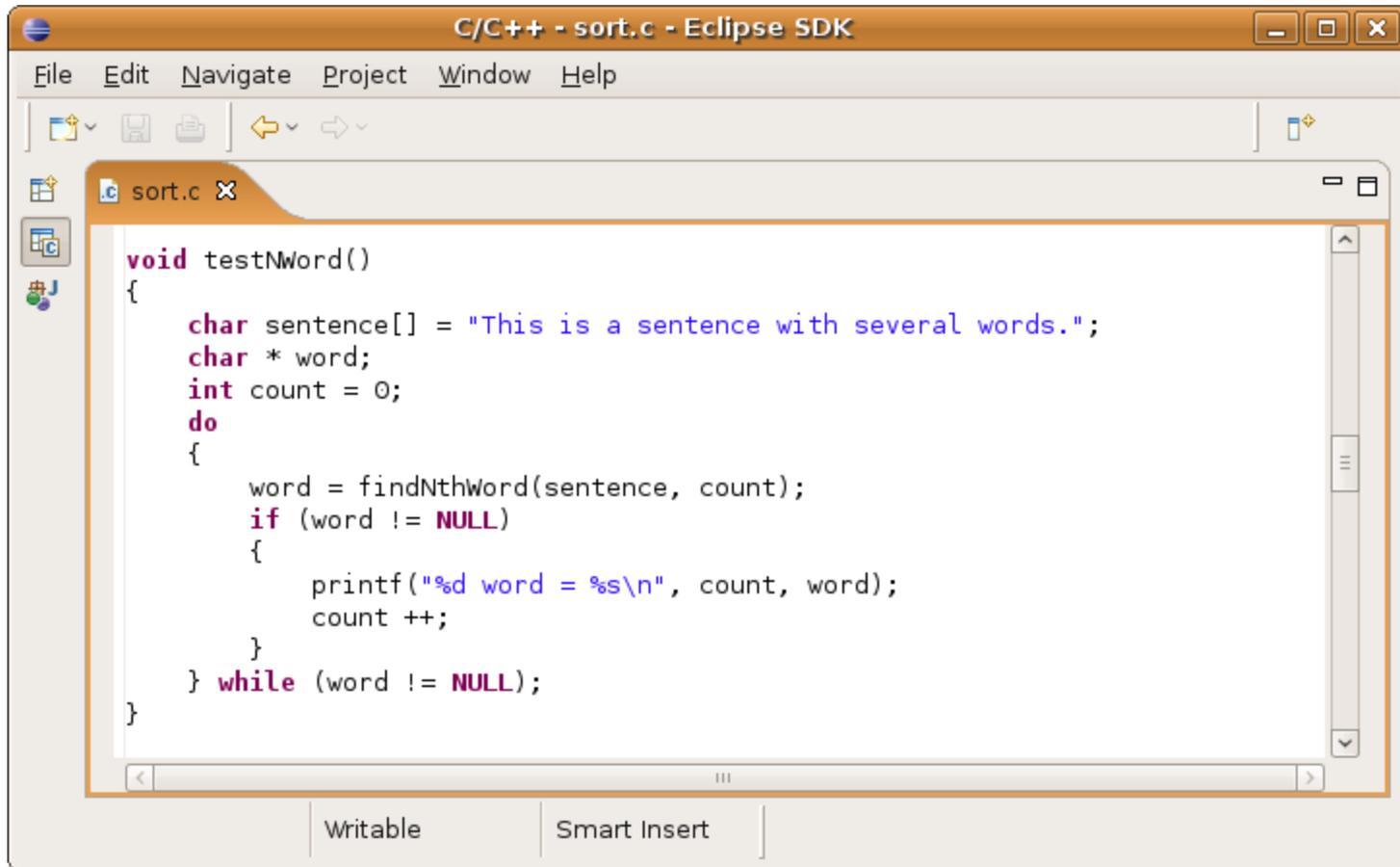
| | | | | | | | | | | | | | | | |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|
| T | h | i | s | | i | s | | a | | w | o | r | d | . | \0 |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|

`wordPtr ++;`

line ↓ **wordPtr** ↓

| | | | | | | | | | | | | | | | |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|
| T | h | l | s | | i | s | | a | | w | o | r | d | . | \0 |
|---|---|---|---|--|---|---|--|---|--|---|---|---|---|---|----|

Test findNthWord

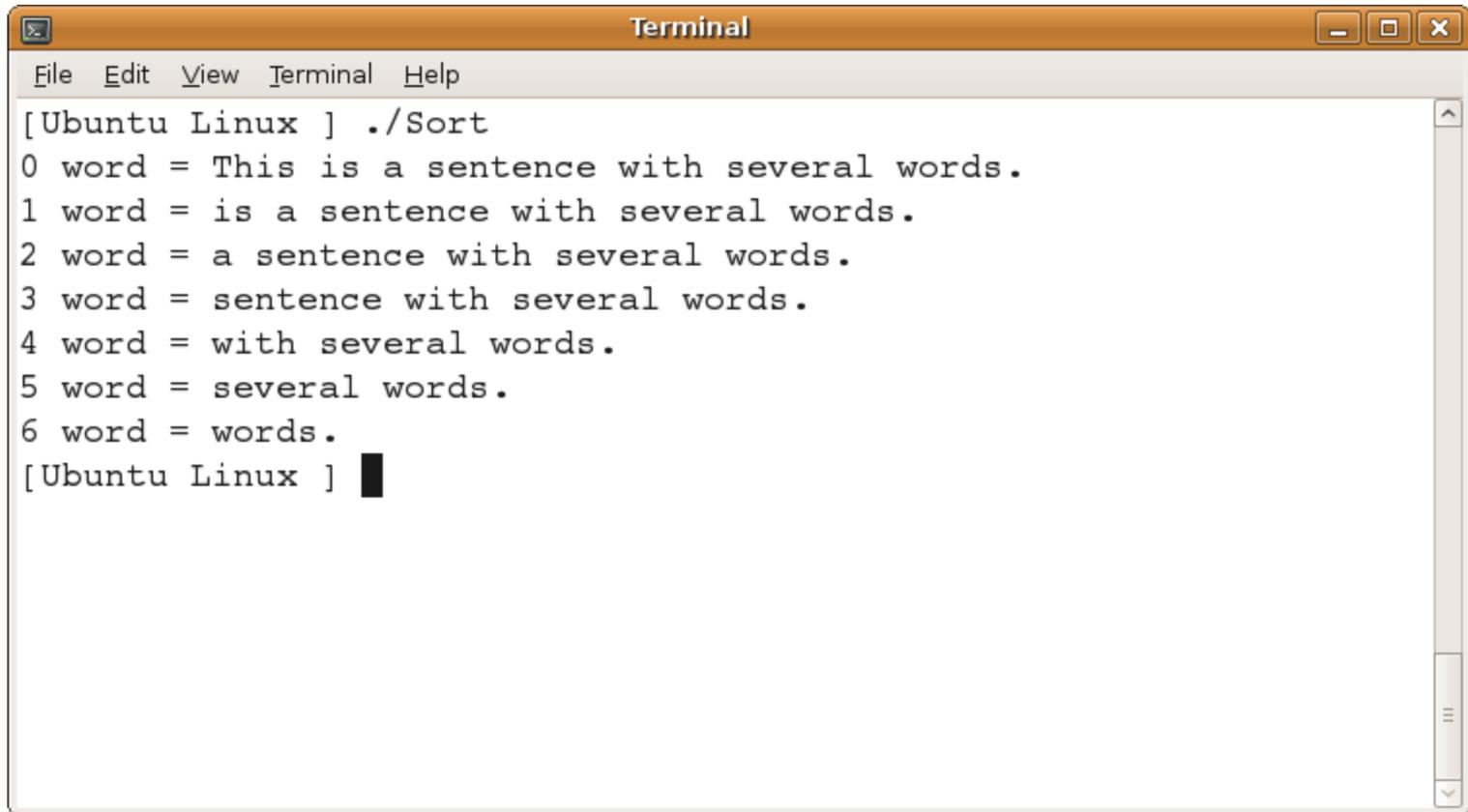


The screenshot shows the Eclipse IDE window titled "C/C++ - sort.c - Eclipse SDK". The menu bar includes File, Edit, Navigate, Project, Window, and Help. The toolbar contains icons for file operations and navigation. The editor window shows the following C code:

```
void testNthWord()
{
    char sentence[] = "This is a sentence with several words.";
    char * word;
    int count = 0;
    do
    {
        word = findNthWord(sentence, count);
        if (word != NULL)
        {
            printf("%d word = %s\n", count, word);
            count ++;
        }
    } while (word != NULL);
}
```

At the bottom of the editor, there are buttons for "Writable" and "Smart Insert".

Test Output

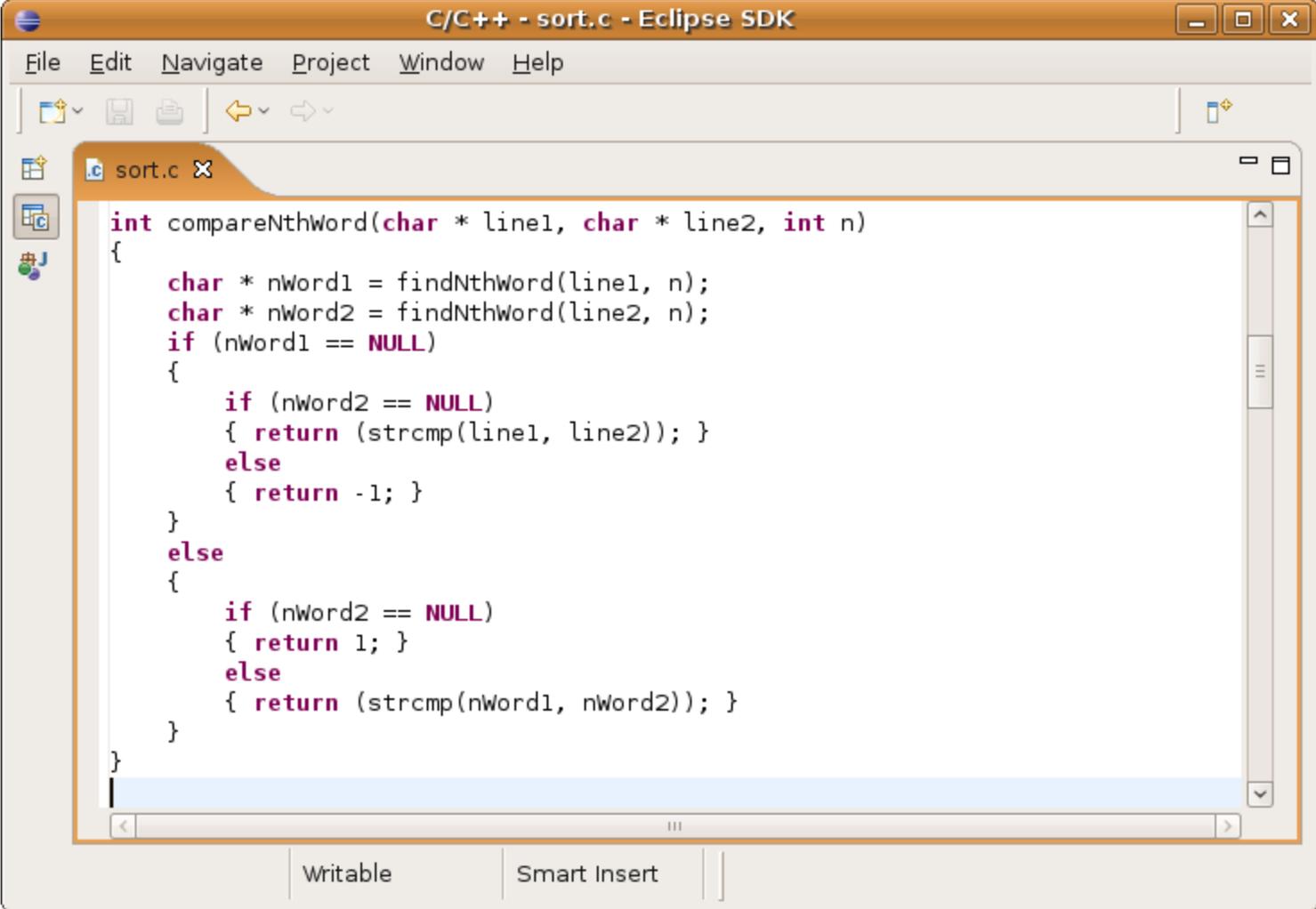


```
Terminal
File Edit View Terminal Help
[Ubuntu Linux ] ./Sort
0 word = This is a sentence with several words.
1 word = is a sentence with several words.
2 word = a sentence with several words.
3 word = sentence with several words.
4 word = with several words.
5 word = several words.
6 word = words.
[Ubuntu Linux ] █
```

Define Line Order

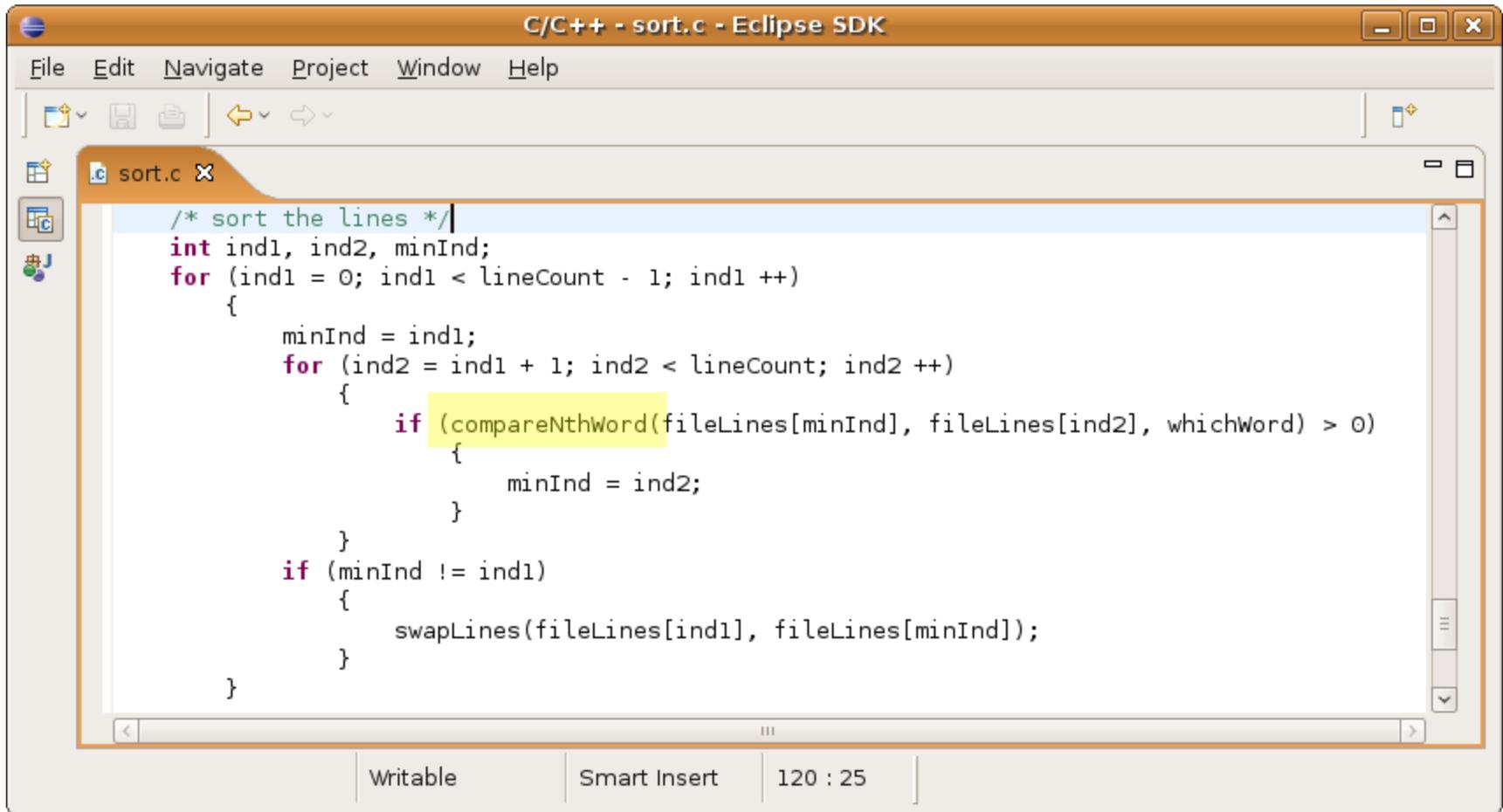
| line1's nth word (word1) | line2's nth word (word2) | order |
|---------------------------------|---------------------------------|----------------------|
| NULL | NULL | strcmp(line1, line2) |
| NULL | not NULL | line1 < line2 (-1) |
| not NULL | NULL | line1 > line2 (1) |
| not NULL | not NULL | strcmp(word1, word2) |

Compare Lines by their n-th Words



```
int compareNthWord(char * line1, char * line2, int n)
{
    char * nWord1 = findNthWord(line1, n);
    char * nWord2 = findNthWord(line2, n);
    if (nWord1 == NULL)
    {
        if (nWord2 == NULL)
        { return (strcmp(line1, line2)); }
        else
        { return -1; }
    }
    else
    {
        if (nWord2 == NULL)
        { return 1; }
        else
        { return (strcmp(nWord1, nWord2)); }
    }
}
```

Sort the Lines by the n-th Words



The screenshot shows the Eclipse IDE interface with a C++ file named 'sort.c'. The code implements a selection sort algorithm based on the n-th word of each line. The 'compareNthWord' function call is highlighted in yellow. The status bar at the bottom indicates 'Writable', 'Smart Insert', and '120 : 25'.

```
C/C++ - sort.c - Eclipse SDK
File Edit Navigate Project Window Help
sort.c
/* sort the lines */
int ind1, ind2, minInd;
for (ind1 = 0; ind1 < lineCount - 1; ind1 ++)
{
    minInd = ind1;
    for (ind2 = ind1 + 1; ind2 < lineCount; ind2 ++)
    {
        if (compareNthWord(fileLines[minInd], fileLines[ind2], whichWord) > 0)
        {
            minInd = ind2;
        }
    }
    if (minInd != ind1)
    {
        swapLines(fileLines[ind1], fileLines[minInd]);
    }
}
```

Writable | Smart Insert | 120 : 25

Summary: What have we learned?

- structure of C programs: main, functions
- data types: int, char, float, array
- flow control: if, while, for, switch, function
- file: write and read
- memory: allocate and release
- algorithm: selection sort