**UNIT 11**

**TEXT EDITORS**

**History**

Be­fore text ed­i­tors ex­isted, com­puter text was punched into [cards](https://wiki2.org/en/Punched_cards) with [key­punch](https://wiki2.org/en/Keypunch) ma­chines. Phys­i­cal boxes of these thin card­board cards were then in­serted into a card-reader. Mag­netic tape and disk "card-im­age" files created from such card decks often had no line-sep­a­ra­tion char­ac­ters at all, and as­sumed fixed-length 80-char­ac­ter records. An al­ter­na­tive to cards was punched paper tape. It could be cre­ated by some [teleprinters](https://wiki2.org/en/Teleprinter), which used spe­cial char­ac­ters to in­di­cate ends of records.

The first text ed­i­tors were ["line ed­i­tors"](https://wiki2.org/en/Line_editor) ori­ented to [type­writer](https://wiki2.org/en/Typewriter)-style ter­mi­nals with­out dis­plays. Com­mands ef­fected edits to a file at an imag­i­nary in­ser­tion point called the "cur­sor". Edits were ver­i­fied by typ­ing a com­mand to print a small sec­tion of the file, and pe­ri­od­i­cally by print­ing the en­tire file. In some line ed­i­tors, the cur­sor could be moved by com­mands that spec­i­fied the line num­ber in the file, text [strings](https://wiki2.org/en/String_(computer_science)) for which to search, and [reg­u­lar expres­sions](https://wiki2.org/en/Regular_expression). Line ed­i­tors were major im­prove­ments over key­punch­ing. Some line ed­i­tors could be used by keypunch.

When [com­puter ter­mi­nals](https://wiki2.org/en/Computer_terminal) with video screens be­came avail­able, [screen-based text ed­i­tors](https://wiki2.org/en/Visual_editor) be­came com­mon. One of the ear­li­est full-screen ed­i­tors was “[O26](https://wiki2.org/en/O26_(text_editor))”, which was writ­ten for the op­er­a­tor con­sole of the [CDC 6000 series](https://wiki2.org/en/CDC_6000_series) com­put­ers in 1967. An­other early full-screen ed­i­tor was “[vi](https://wiki2.org/en/Vi)”. Writ­ten in the 1970s, it is still a stan­dard editor on [Unix](https://wiki2.org/en/Unix) and [Linux](https://wiki2.org/en/Linux) op­er­at­ing sys­tems.

Also writ­ten in the 1970s was the [UCSD Pas­cal](https://wiki2.org/en/UCSD_Pascal) Screen Ori­ented Ed­i­tor, which was op­ti­mized both for indented source code as well as gen­eral text.

[Emacs](https://wiki2.org/en/Emacs), one of the first [free and open source soft­ware](https://wiki2.org/en/Free_and_open_source_software) pro­jects, is an­other early full-screen or real-time ed­i­tor, one that was [ported](https://wiki2.org/en/Port_(software)) to many systems. A full-screen ed­i­tor's ease-of-use and speed (com­pared to the line-based editors) mo­ti­vated many early pur­chases of video terminals.

The core data struc­ture in a text ed­i­tor is the one that man­ages the se­quence of char­ac­ters or list of [records](https://wiki2.org/en/Storage_record) that rep­re­sents the cur­rent state of the file being edited. While the for­mer could be stored in a sin­gle long con­sec­u­tive [array](https://wiki2.org/en/Array_data_structure) of char­ac­ters, the de­sire for text ed­i­tors that could more quickly in­sert and delete text.

**Vocabulary:**

punch – вбивать

keypunch machine – клавишный перфоратор

physical box –

cardboard card – перфокарта

insert – вставлять

card-reader – устройство считывания с перфокарт

magnetic tape – магнитная лента

card-image file – отображение данных перфокарты

card deck – колода перфокарт

line-separation character – разделительный символ

assume – принимать

punched paper-tape – бумажная перфолента

teleprinter – телетайп

indicate – отображать

line editor – редактор строк

[type­writer](https://wiki2.org/en/Typewriter)-style ter­mi­nal – клавиатура по типу пишущей машинки

edits – [система передачи изображений](https://translate.academic.ru/%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0%20%D0%BF%D0%B5%D1%80%D0%B5%D0%B4%D0%B0%D1%87%D0%B8%20%D0%B8%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%B8%D0%B9%20%D1%82%D0%B5%D1%85%D0%BD%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D1%85%20%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%BE%D0%B2/en/ru/)

insertion point – место ввода

text string – текстовая строка

regular expression – регулярное выражение

[screen-based text ed­i­tors](https://wiki2.org/en/Visual_editor) – экранный текстовый редактор

op­er­a­tor con­sole – пульт оператора

in­dented source code – закрытый исходный текст

source soft­ware project – программа с открытыми исходными кодами

ease-of-use – простой в управлении

pur­chase – покупка

core data struc­ture – структура основных данных

se­quence of char­ac­ters – порядок символов

cur­rent state – текущее состояние

con­sec­u­tive [array](https://wiki2.org/en/Array_data_structure) – последовательный порядок

1. **Match the following words to make word combinations:**
2. source
3. punched
4. cardboard
5. consecutive
6. computer
7. real-time
8. line
9. keypunch
10. sequence of
11. software
12. screen-based
13. operating
14. insertion
15. card
16. line
17. text
18. video
19. operator
20. systems
21. paper tape
22. editors
23. array
24. cards
25. characters
26. strings
27. code
28. decks
29. terminals
30. projects
31. terminals
32. console
33. text editors
34. point
35. number
36. editor
37. machines
38. **Answer the questions:**
39. Where was computer text punched before the invention of text editors?
40. What does the core data structure manage in a text editor?
41. What is a standard editor on Unix?
42. How did keypunch machines work?
43. What was an alternative to cardboard cards?
44. What were line editors oriented to?
45. How were edits verified?
46. What is one of the first free source software projects?
47. When did screen-based text editors become common?

**Types of texts**

There are important differences between plain text (created and edited by text editors) and [rich text](https://wiki2.org/en/Rich_text) (created by [word processors](https://wiki2.org/en/Word_processor) or [desktop publishing software](https://wiki2.org/en/Desktop_publishing_software)).

Plain text consists of character representation. Each character is represented by a fixed-length sequence of one, two, or four bytes, or as a variable-length sequence of one to four bytes, in accordance to specific [character encoding](https://wiki2.org/en/Character_encoding) conventions.

These conventions define many printable characters, but also [non-printing characters](https://wiki2.org/en/Whitespace_character) that control the flow of the text, such [space](https://wiki2.org/en/Space_(punctuation)), [line break](https://wiki2.org/en/Newline), and [page break](https://wiki2.org/en/Page_break). Plain text contains no other information about the text itself, not even the character encoding convention employed. Plain text is stored in [text files](https://wiki2.org/en/Text_file), although text files do not exclusively store plain text.

Rich text, on the other hand, may contain metadata, character formatting data, paragraph formatting data, and page specification data. Rich text can be very complex. Rich text can be saved in [binary format](https://wiki2.org/en/Binary_format), text files adhering to a [markup language](https://wiki2.org/en/Markup_language), or in a hybrid form of both.

Text editors are intended to open and save text files containing either plain text or anything that can be interpreted as plain text, including the markup for rich text.

Text editors have 5 main features:

[Find and replace](https://wiki2.org/en/Find_and_replace). Text editors provide extensive facilities for searching and replacing text, either on groups of files or interactively. Advanced editors can use [regular expressions](https://wiki2.org/en/Regular_expression) to search and edit text or code.

[Cut, copy, and paste](https://wiki2.org/en/Cut,_copy,_and_paste). Most text editors provide methods to duplicate and move text within the file, or between files.

Ability to handle [UTF-8](https://wiki2.org/en/UTF-8) encoded text.

[Text formatting](https://wiki2.org/en/Text_formatting). Text editors often provide basic visual formatting features like [line wrap](https://wiki2.org/en/Line_wrap), [auto-indentation](https://wiki2.org/en/Indentation_(typesetting)), [bullet list](https://wiki2.org/en/Bullet_list) formatting using [ASCII](https://wiki2.org/en/ASCII) characters, [comment formatting](https://wiki2.org/en/Comment_(computer_programming)), [syntax highlighting](https://wiki2.org/en/Syntax_highlighting) and so on. These are typically only for display and do not insert formatting codes into the file itself.

[Undo and redo](https://wiki2.org/en/Undo). Text editors provide a way to undo and redo the last edit, or more.

**Vocabulary:**

plain text – обычный текст

rich text – форматированный текст

word processor – текстовый процессор

desktop publishing software – программное обеспечение для настольных издательских систем

character – символов

fixed-length sequence – фиксированная последовательность

variable-length sequence – переменная последовательность

character encoding – кодирование символов

convention – соглашение

the flow of the text – поток данных

line break – разрыв строки

page break – разрыв страницы

employed – применяемый

store – хранить

exclusively – исключительно

metadata – метаданные

character formatting data – данные форматирования символа

paragraph – абзац

page specification data – данные о странице

complex – сложный

binary – двоичный

adhering – придерживающийся

markup language – язык разметки

provide – предоставлять

facility – возможность

interactively – интерактивный режим

advanced – расширенный

expression – выражение

paste – вставка

within – внутри

handle – обработка

feature – функция

line wrap – перенос строк

auto-indentation – автоматическое отступление

bullet list – список маркеров

syntax highlighting – выделение синтаксических конструкций

display – отображать

undo – отмена

redo – восстановление

1. **Answer the questions:**
2. What is the difference between plain text and rich text?
3. What do non-printing characters control?
4. What are the typical creatures of text editors?
5. What are the basic visual formatting features?
6. What does rich text contain?
7. What does plain text contain?
8. What way can rich text be saved?
9. **Match the following words:**
10. printable
11. text
12. syntax
13. line
14. fixed-length
15. formatting
16. character
17. character
18. line
19. bullet
20. markup
21. binary
22. plain
23. comment
24. paragraph
25. word
26. page
27. last
28. formatting data
29. editor
30. processors
31. highlighting
32. break
33. representation
34. formatting
35. edit
36. characters
37. list formatting
38. format
39. codes
40. encoding
41. sequence
42. language
43. specification
44. wrap

text