

Задача за корекция на правописа

Задание:

Да се разработи програма, която коригира българския правопис в зададен текст по зададени правила - правилата, което трябва да се реализират са дадени в секция „Шаблони за корекция“. Името на файла с текста за кодиране и корекция и името на изходен файл се задават на командния ред.

Всеки коригиран параграф се изписва на екрана по следния ред:

- Първо се извежда на екрана оригиналния текст, като с червен цвят на бял фон се показват символите, които се премахват в крайния вариант. След това се извежда поправения текст, като с жълт цвят на син фон се отбелязват добавените символи.
- Извеждането на всеки следващ параграф да започва с натискане на клавиша „Enter“.
- За оцветяване на текста на екрана да се използват ANSI Escape кодове.

В изходния файл се извежда само редактирания текст без информация за местата на направените корекции.

Речник

[<>]+ -> поне един екземпляр от вида в <>

<печатим символ> -> букви, цифри

<гб> -> голяма буква

<мб> -> малка буква

<пз> -> ,:<ки> // препинателен знак

<ки> -> .! ? // край изречение

<нр> -> нов ред

<таб> -> табулация (хоризонтална)

<шп> -> интервал (шпация)

<дума> -> последователност от една или повече букви или цифри

<ндгр1> -> Например, Всъщност, Според мен, Наистина, Следователно, Действително, Реално, Практически, Така, Може би, Значи, Вероятно, Очевидно, Навярно, По всяка вероятност, В действителност, В края на краищата

<ндгр2> -> Казват, Напротив, Разбира се, Изглежда, Струва ми се, Види се, Общо взето, Тъй да се каже, Между нас казано, Както казват, Първо, Второ, Трето, Естествено

<ндгр3> -> казват, напред, разбира се, изглежда, струва ми се, види се, общо взето, тъй да се каже, нежду нас казано, както казват

<ндгр4> -> а, но, че, и, не, или

<ндгр5> -> дори, само, именно, едва, особено, поне, тъкмо, под, много, малко, даже, чак, щом, тъй, след, освен, чак, единствено, именно, тъкмо

- <ндгр6>** -> който, която, което, които, чийто, чиято, чието, чиито
- <ндгр7>** -> за, в, с, заради, от, по, поради, вследствие на, благодарение на
- <ндгр8>** -> а, ала, но, макар и, ако и, та, ама, пък, що, дето, обаче, ами, защото, отколкото, без да, докато, като
- <ндгр9>** -> само, дори, едва, чак, даже, единствено, именно, тъкмо, много, малко, не, и
- <ндгр10>** -> за да, без да, като да, освен да, въпреки че, освен че, само че, тъй че, така че, макар че, сякаш че, затова че, задето, освен дето, освен като, след като, макар да, ако и да, когато, ако, дето, и, или, че, да
- <ндгр11>** -> макар, въпреки, така, само, при положение, в случай, по причина, при условие, при все, сякаш, освен, независимо, Въпреки, Така, Само, При положение че, Макар, В случай че, При условие, При все, Освен, Независимо
- <ндгр12>** -> щом, тъй, след, освен
- <ндгр13>** -> както следва, вместо да, за разлика от
- <ндгр14>** -> а, но, че, и

Грешка	Корекция
--------	----------

Шаблони за корекция

1. Премахване на интервал между дума и многоточие

<дума><шп>...	<дума>...
---------------	-----------

2. Премахвана не излишна точка

<дума>..	<дума>.
<дума>...[.]+	<дума>... // повече от три точки се свеждат до 3 точки

3. Корекция на неправилно пренесен ред при пряка реч

<дума>,<шп><нр> -<шп><дума>	<дума>,- <шп><дума>
<дума>,<шп><нр> -<дума>	<дума>,- <шп><дума>
<дума>,<шп><нр> <шп>-<шп><дума>	<дума>,- <шп><дума>
<дума>,<шп><нр> <шп>-<дума>	<дума>,- <шп><дума>

4. Премахване на интервали между препинателни знаци

<пз><шп><пз>	<пз><пз>
--------------	----------

5. Премахване на интервал пред препинателен знак.

<шп><пз>	<пз>
----------	------

6. Премахване на ',' от препинателен знак

<пз> ,	<пз>
<пз> ,	<пз>
,<пз>	<пз>
,<пз>	<пз>
ако има повече от 1 интервал, то всичките се премахват заедно със ','	

7. Добавяне на интервал между ',' и печатим символ

,<дума>	,<шп><дума>
---------	-------------

8. Премахване на интервали между препинателни знаци

<пз><шп><пз>	<пз><пз>	
<пз><шп>[<шп>]+<пз>	<пз><пз>	// две или повече шпации

9. Премахване на ',' след начална дума в изречението

<ндгр1> ,	<ндгр1><шп>
<ндгр1>,<шп>	<ндгр1><шп>

10. Задължителна ',' след начална дума в изречението

<ндгр2><шп>	<ндгр2>,<шп>
-------------	--------------

11. Задължителна ',' пред и след дума

<дума><шп><ндгр3><шп>	<дума>,<шп><ндгр3>,<шп>
<дума><шп><ндгр3> ,	<дума>,<шп><ндгр3> ,
,<шп><ндгр3><шп>	,<шп><ндгр3>,<шп>
,<ндгр3> ,	,<шп><ндгр3> ,

12. Премахване на ',' след съюз

<ндгр4><шп> ,	<ндгр4><шп>
---------------	-------------

<ндгр4>,	<ндгр4><шп>
----------	-------------

13. Преместване на ','

<ндгр5><шп>,<ндгр5><шп>	<ндгр5><шп>
<ндгр5>,<ндгр5><шп>	<ндгр5><шп>

14. Премахване на ',' след уточняващо наречие

<ндгр7><шп>,<ндгр6>	<ндгр7>,<шп><ндгр6>
<ндгр7>,<ндгр6>	<ндгр7>,<шп><ндгр6>

15. Добавяне на ','

не,<шп>защото	не защото
не,защото	не защото
<ндгр5>,<шп>защото	<ндгр5> защото
<ндгр5>,защото	<ндгр5> защото

16. Добавяне на ','

За съжаление <дума>	За съжаление, <дума>
<дума> за съжаление <дума>	<дума>, за съжаление, <дума>
<дума>, за съжаление <дума>	<дума>, за съжаление, <дума>
<дума>,за съжаление <дума>	<дума>, за съжаление, <дума>
<дума> за съжаление, <дума>	<дума>, за съжаление, <дума>
<дума> за съжаление,<дума>	<дума>, за съжаление, <дума>

17. Премахване на ','

<ндгр9><шп>,<ндгр10>	<ндгр9><шп><ндгр10>
<ндгр9>,<ндгр10>	<ндгр9><шп><ндгр10>

18. Премахване на ',' пред 'че'

<ндгр11><шп>,че	<ндгр11><шп>че
<ндгр11>,че	<ндгр11><шп>че

19. Премахване на ','

че,<шп>за да	че за да
че,за да	че за да

20. Добавяне на ','

<ндгр12><шп>като	, <ндгр12><шп>като
------------------	--------------------

21. Добавяне на ','

<дума><шп><ндгр13>	<дума>, <шп><ндгр13>
--------------------	----------------------

22. Премахване на ',' пред 'естествено'

<ндгр13> , естествено	<ндгр13> естествено
<ндгр13>, естествено	<ндгр13> естествено

23. Кога се слага 'във'

<шп>в в	<шп>във в
<шп>в ф	<шп>във ф

24. Кога се слага 'със'

<шп>с с	<шп>със с
<шп>с з	<шп>със з

25. Добавяне на ','

<дума><шп><ндгр8>	<дума>, <шп><ндгр8>
-------------------	---------------------