

Aufgabe (1)

[4 Punkte]

Definieren Sie für die Klasse **Document** eine Methode bzw. ein Prädikat **wort_normalerweise_klein?**, das ein Token (String) als Argument nimmt und den Wert *true* zurückgibt, wenn das Token im Dokument in der Regel mit einem initialen Kleinbuchstaben beginnt. Wenn also z.B. der Artikel |der| im Dokument 4x klein geschrieben und 1x groß geschrieben vorkommt, liefert die Methode den Wert *true*.

Aufgabe (2)

[6 Punkte]

Definieren Sie für die Klasse **Document** eine Methode **klassifiziere_punkte**, die die Token im Dokument, die mit einem Punkt enden, untersucht, um festzustellen, ob es bei dem Punkt um einen Abkürzungspunkt oder ein Satzendezeichen handelt. (String) als Argument nimmt und anhand des letzten Zeichens des Satzes Kriterien:

1. Wenn das nächste Token mit einem Kleinbuchstaben beginnt, liegt eine Abkürzung vor.
2. Wenn das nächste Token mit einem Großbuchstaben beginnt, aber sonst in der Regel im Dokument klein geschrieben wird, liegt ein Satzendezeichen vor.
3. Wenn das Token interne Punkte enthält, liegt eine Abkürzung vor.
4. Sonst ist es eine Satzendezeichen.

Aufgabe (3)

[4 Punkte]

Legen Sie eine Testdatei mit sinnvollen Tests für diese Aufgaben an.