Computerlinguistische Programmiertechnik Beliankou • Beyer • Naumann WS 2016

Aufgabenblatt (1)

Aufgabe (1) [2 Punkte]

Weisen Sie einen kurzen Text (z.B eine Spiegel-Online-Nachricht) einer Variable zu und definieren Sie anschließend die Funktion *token-liste*, die einen Text (string) als Argument nimmt und eine Liste aller im Text enthaltenen Token als Wert liefert.

Beispiel

```
> (token-liste "Aller Anfang ist schwer")

⇒ ("Aller" "Anfang" "ist" "schwer")
```

Aufgabe (2) [3 Punkte]

Definieren Sie die Funktion *trim-token*, die ein Token (string) als Argument nimmt und als Wert eine Kopie des Token liefert, bei dem alle Sonderzeichen am Anfang bzw. Ende des Tokens entfernt wurden:

Beispiel

```
> (trim-token "Anfang,")

⇒ "Anfang")

> (trim-token "(Anfang)!")

⇒ "Anfang")
```

Aufgabe (3) [4 Punkte]

Definieren Sie die Funktion frequenzliste, die eine Tokenliste als Argument nimmt und als Wert eine Frequenzliste (map) erzeugt, die alle gesäuberten (trim-token) Token enthält, die aus mindestens zwei Zeichen bestehen.

Beispiel

```
> (frequenzliste
"aller, aller Anfang ist schwer.") \Rightarrow {"aller" 2, "Anfang" 1, "ist" 1, "schwer" 1}
```

Aufgabe (4) [2 Punkte]

Definieren Sie die Funktion hapaxlegomena, die eine Frequenzliste (map) als Argument nimmt und als Wert eine Liste aller Hapaxlegomena aus dieser Liste liefert.

Beispiel

```
> (hapaxlegomena {"aller" 2, "Anfang" 1, "ist" 1, "schwer" 1}) \Rightarrow ("Anfang" "ist" "schwer")
```