Problem 1. This exercise traces through the first few steps of a perceptron training algorithm. The task is to classify a sentence into one of two classes, which are called $+1$ and -1. We will use just two features; unlike the example in class, these features are not binary, but integer-valued. The features are:

pron The number of personal pronouns in the sentence.

noun The number of proper and common nouns in the sentence.

In the data below, each instance of **pron** is marked in **red boldface**, and each instance of *noun* is marked in *green bold italics*. A hyphenated term such as *Cochin-China* or *great-aunt* is considered a single term. The data (classes and sentences) are taken from Argamon et al.: Gender, genre, and writing style in formal written texts, *Text* **23**(3): 321–346, 2003.

a. (15 points) Count the features in each sentence and update the perceptron weights, using Algorithm 5 of the perceptron reading (Daumé).

$+1$ *Clara* never failed to be astonished by the extraordinary *felicity* of **her** own *name*.

Feature counts: **pron** __1__ *noun* __3__   Weights: **pron** __1__ *noun* __3__ bias __1__

-1 By 1925 present-day *Vietnam* was divided into three *parts* under French colonial *rule*.

Feature counts: **pron** __0__ *noun* __3__   Weights: **pron** __1__ *noun* __0__ bias __0__

$+1$ **She** found it hard to trust **herself** to the mercy of *fate*, which had managed over the *years* to convert **her** greatest *shame* into one of **her** greatest *assets*, and even after *years* of comparative *security* **she** was still prepared for, still half expecting the old *gibes* to be revived.

Feature counts: **pron** __5__ *noun* __7__   Weights: **pron** __1__ *noun* __0__ bias __0__

-1 The southern *region* embracing *Saigon* and the *Mekong delta* was the *colony* of *Cochin-China*; the central *area* with its imperial *capital* at *Hue* was the *protectorate* of *Annam*; and the northern *region*, *Tongking*, was also a separate *protectorate* with its *capital* at *Hanoi*.

Feature counts: **pron** __0__ *noun* __16__   Weights: **pron** __1__ *noun* __-16__ bias __-1__

$+1$ But whenever **she** was introduced, nothing greeted the amazing, all-revealing *Clara* but *cries* of "How delightful, how charming, how unusual, how fortunate," and **she** could foresee a *time* when *friends* would name **their** *babies* after **her** and refer back to **her** with *pride* as the *original* from which *inspiration* had first been drawn.

Feature counts: **pron** __5__ *noun* __8__   Weights: **pron** __6__ *noun* __-8__ bias __0__

-1 The Annamese *emperor*, *Khai Dinh*, in *theory* ruled the two northern *regions* from *Hue* with the *benefit* of French *protection*, while *Cochin-China* was governed directly from *Paris* but in *effect* all three *territories* were ruled as *colonies*.

Feature counts: **pron** __0__ *noun* __13__   Weights: **pron** __6__ *noun* __-8__ bias __0__