Paper 1: Naive Bayes

link: http://www.aclweb.org/anthology/W98-1119
implementation example: https://nlp.stanford.edu/courses/cs224n/2006/fp/report-hicks-wright.pdf

In this paper, we learn about a naive Bayes method for pronoun coreference for anaphoras.

What is Naive Bayes?

Naive Bayes is a method using Bayesian theorem (probability of A given B happens) with a simple (naive) assumption that features are independent. It delivers us probabilities of an outcome.

What are anaphoras?

Anaphoras are when object mention comes before its pronoun. Last sentence contains an example of an anaphora - we have „object mention", and then „its".

So, what's the method?

We extract some data from the text and then apply a naive Bayes classifier.

What are the data?

There are few features extracted from the text:

* distance between a subject and the pronoun
* syntatic situation, or where the subject may be. We can use a breadth-first search of a parse tree (recursing into other sentences when needed), left-to-right (to prioritise anaphoras), with use of restrictions on how pronouns can be used (particularly in case of reflexive pronouns, but these don't seem to be used for this task - still, worthy of looking for limitations with other types of pronouns) - this can be done with so-called Hobbs' method.
* proposed subject: what is its gender, the number (i.e. one or many), animaticity (object or a person) and maybe selectional restriction (verbs that use a pronoun limit what the subject may be, e.g. you usually can't eat a TV set)
* mention count, or how many times a given subject has been mentioned in a text before the pronoun

What is the formula?

Basically, probability of a subject being correct given all above occur. There is a formula for it, but…

$$
\begin{aligned}
F(\rho) \\
&= \arg\max_{a} P(A(\rho) = a | \rho, h, \vec{W}, t, l, s_\rho, \vec{d}, \vec{M}) \\
&= \arg\max_{w_a} P(d_H | a) P(\rho | w_a) \\
&\qquad \frac{P(w_a | h, t, l)}{P(w_a | t, l)} P(a | m_a)
\end{aligned}
$$

let's just say it isn't very readable.

Does it work well?

Well, a bit. 70.3% accuracy rate, but compared to 65.3% accuracy of Hobbs method alone doesn't seem that great. However, because mention count matters, this method works great on texts where the proper subject of pronoun appears more than once.

Can it be improved?

Yes, as shown by implementation example, where students have used two methods: improving pronoun-antecedent classification with maximum entropy classifier on n-grams and by using a language model to gauge how „natural" a sentence would sound if we replaced the pronoun. But interestingly enough, that has led to **worse** results than with Hobb's alone. In their opinion, combining improvement methods has caused these problems, so that can somehow be improved on.