

Decision Tree Development

Akshay Jaggi

March 25, 2018

Main Text

To demonstrate a potential methodology for discerning the use of connectives in context, we modeled connective learning using a random forest approach. Random forests are an ensemble model that aggregates many decision trees. A decision tree is a series of binary decisions that classifies data into two or more classes. The order and structure of these decisions can be learned from data. Here, we randomly subsample subsets of the annotated CHILDES corpus as training data and then test on the remaining data.

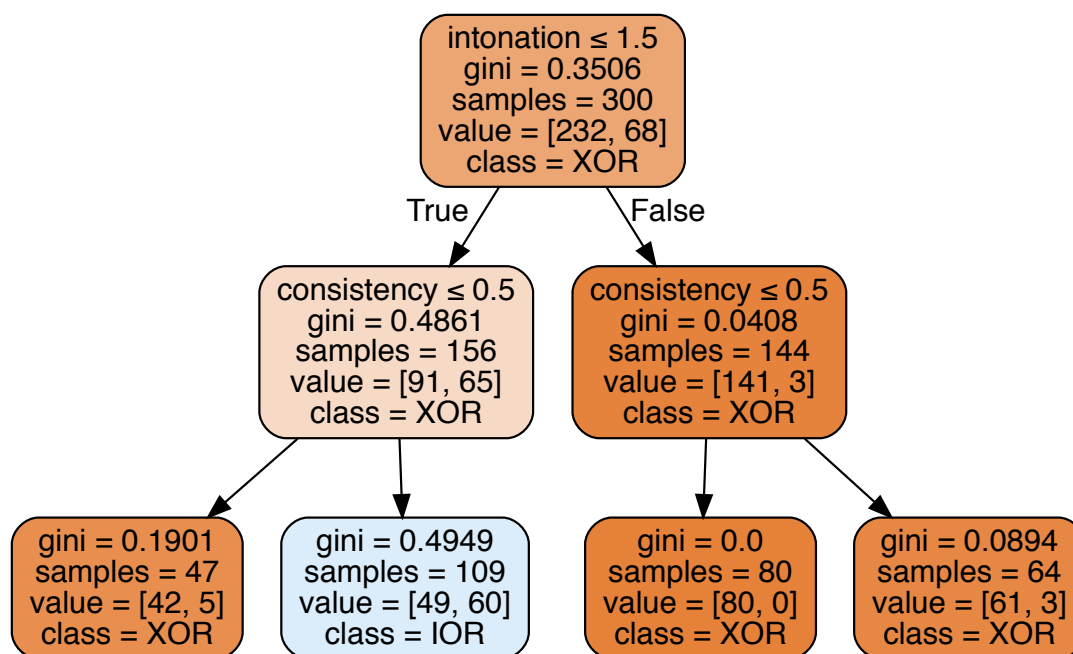
Discerning XOR-IOR

From 300 annotated examples labeled as either IOR or XOR, we use intonation and consistency to separate the two categories.

If intonation is rise-fall, XOR.

If inconsistent, XOR.

Otherwise, IOR



Discerning XOR-IOR-AND

From 300 annotated examples labeled IOR, XOR, or AND, we use intonation, speech act, consistency, and annotation to separate them.

If Annotation is AND, AND

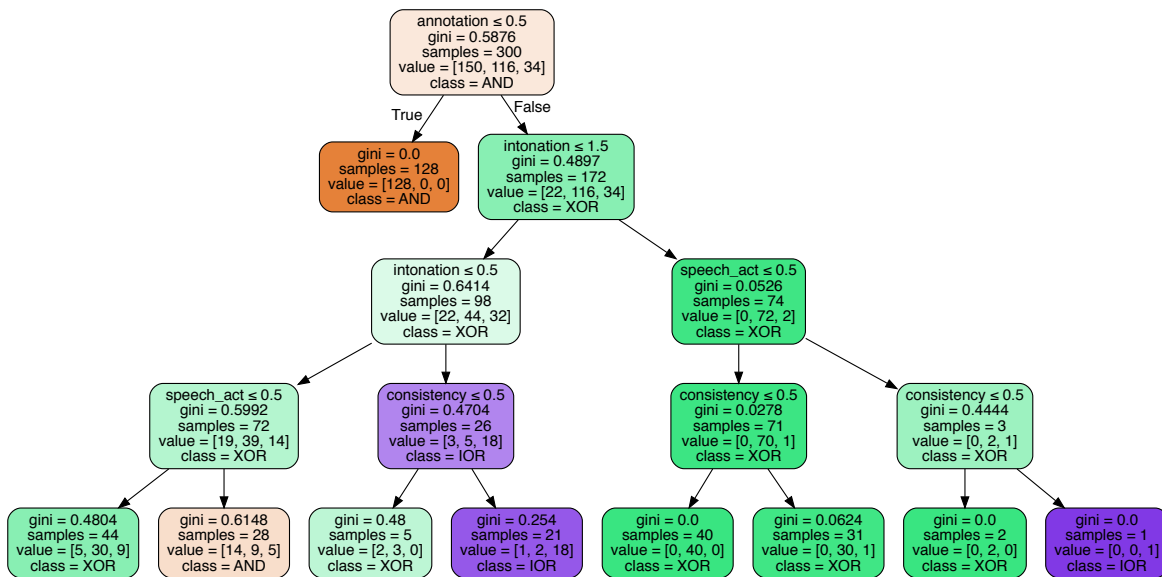
If intonation is rise-fall, XOR

If intonation is rising and inconsistent, XOR

If intonation is rising and consistent, IOR

If intonation is flat and speech act options/defex/unconditional, AND

If intonation is flat and other speech act, XOR



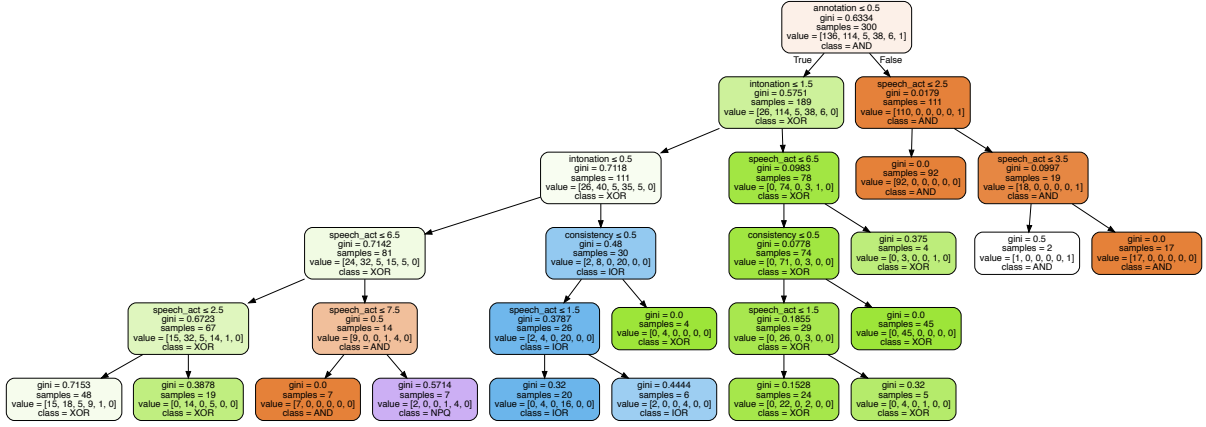


Figure 1: “Everything Tree”

Discerning Everything

From 300 annotated examples labeled IOR, XOR, AND, NPQ, XOR, or NOR, we use intonation, speech act, consistency, utterance type, and annotation to separate them.

If Annotation is AND, AND

If intonation is rise-fall, XOR

If intonation is rising and inconsistent, XOR

If intonation is rising and consistent, IOR

If intonation is flat and speech act is (unconditional or repair or directive description), NPQ

If intonation is flat and speech act is defex, AND

If intonation is flat and speech act is (conditional, identification, clarification, or directive), XOR

If intonation is flat and speech act is (description, options, preference), ???