

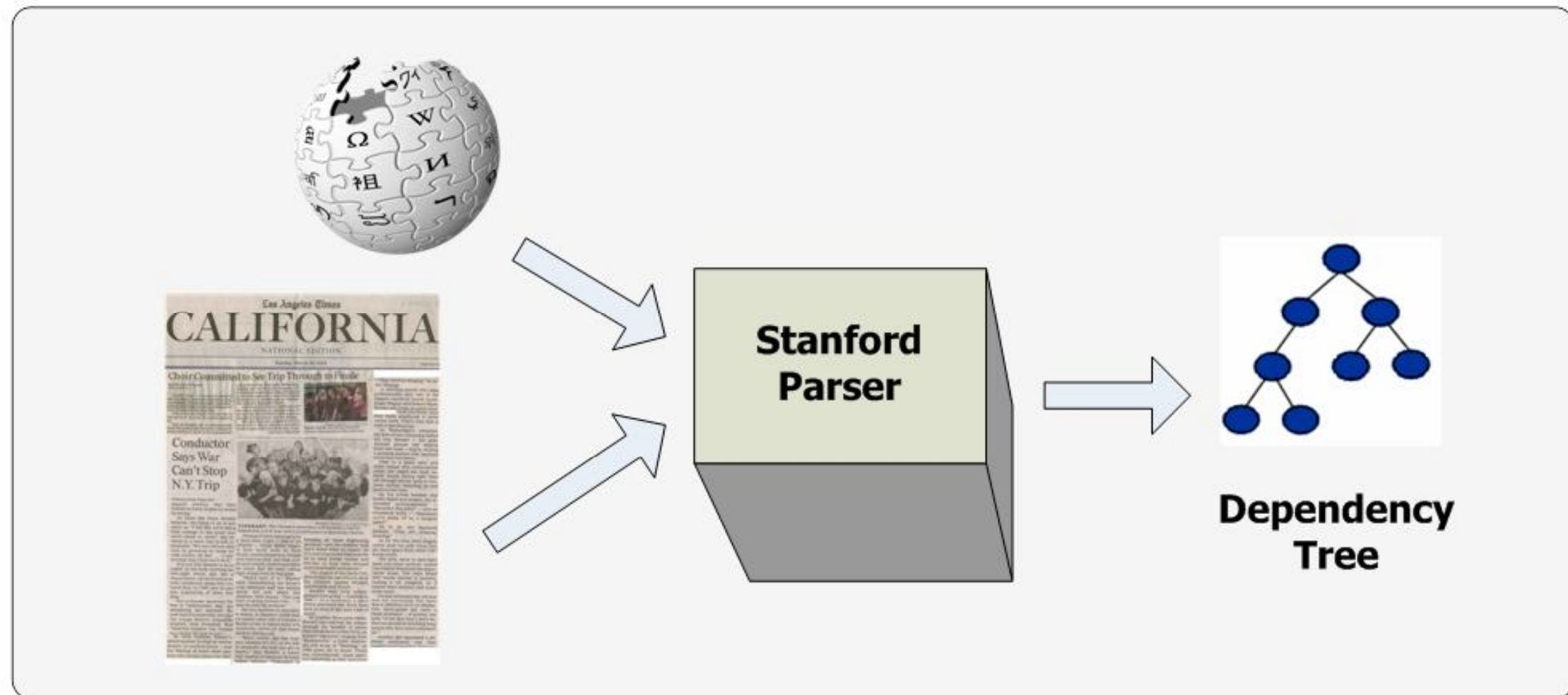
Stanford Parser

DTGenerator

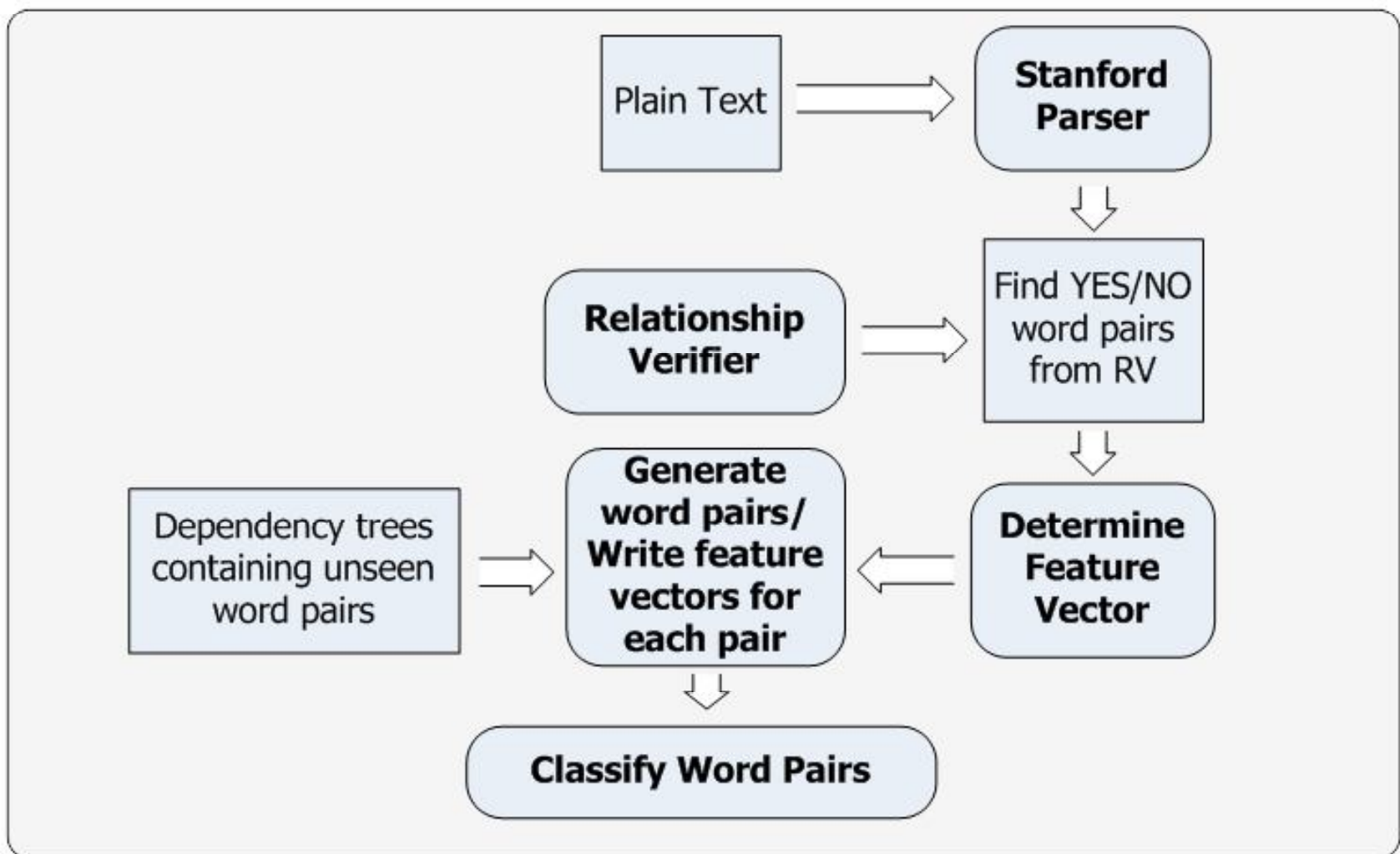
- Parses text using the Stanford Parser
- Loads/preprocesses dependency tree collections

DT

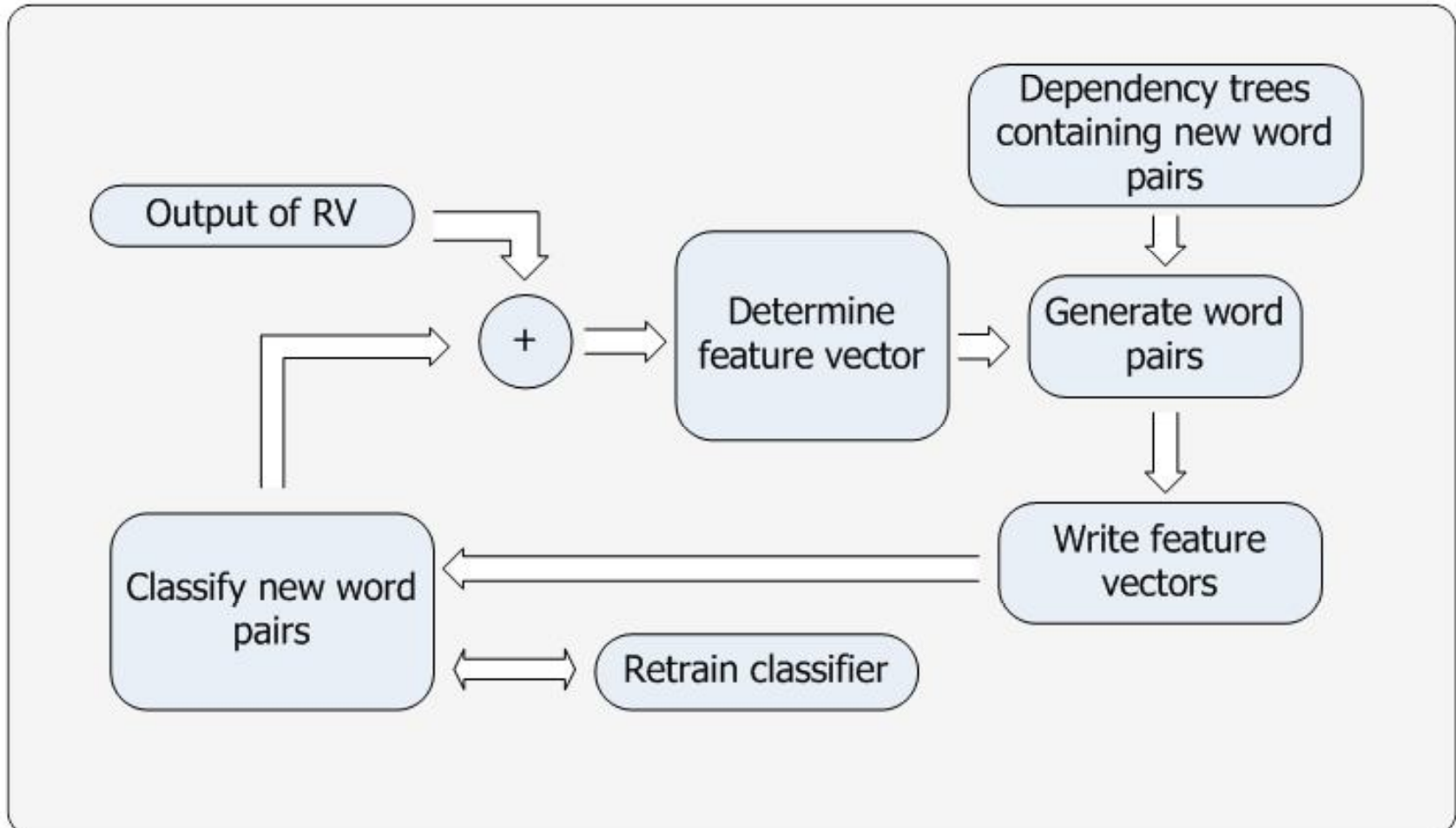
- Holds a single dependency tree
- Creates paths between nouns in the tree (collecting relevant satellites)
- Caches results for speed



System Diagram



Feedback Loop



Relation Verifier

WordNetRV

- Generates pairs with relation
- Prunes pairs to set that exist in corpus
- Generates pairs without relation from corpus (with verification by WordNet)
- Saves results to disk. Subsequent executions load results from files.

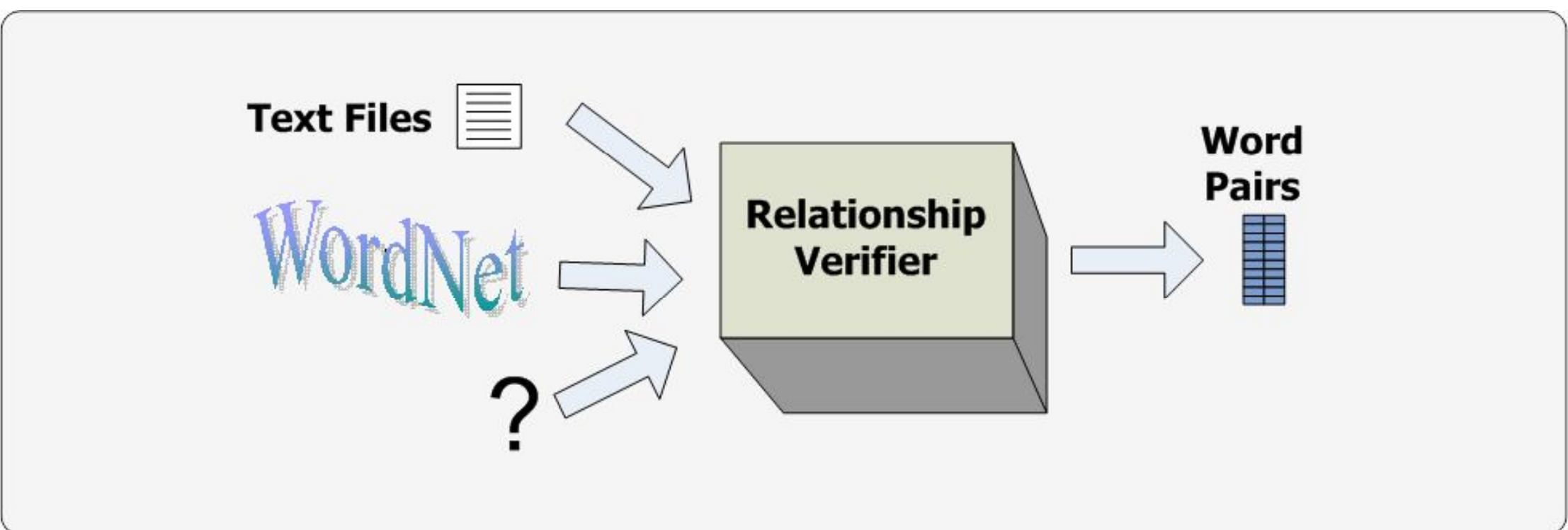
TextFileRV

- Uses user-provided text files instead of WordNet

Supported Relations

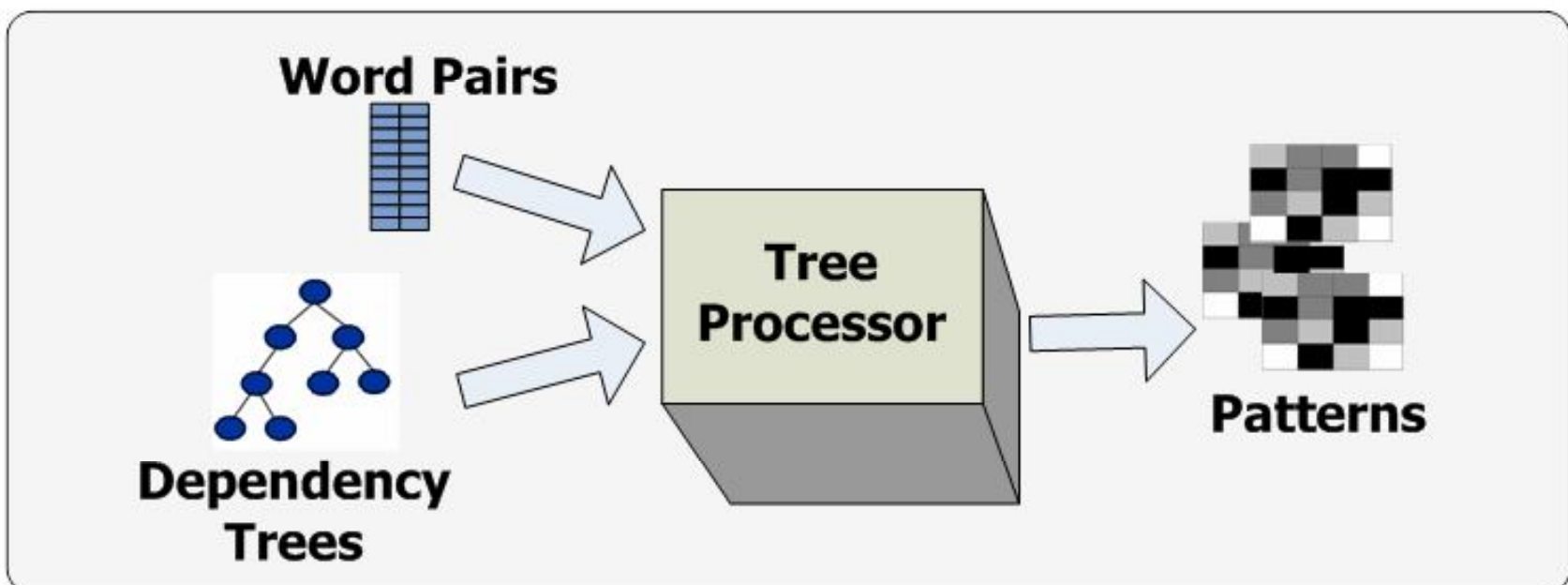
- Antonyms
- Holonyms
- Hypernyms
- Hyponyms
- Meronyms
- Participles
- Synonyms

Custom RVs can also be created and plugged into the framework.



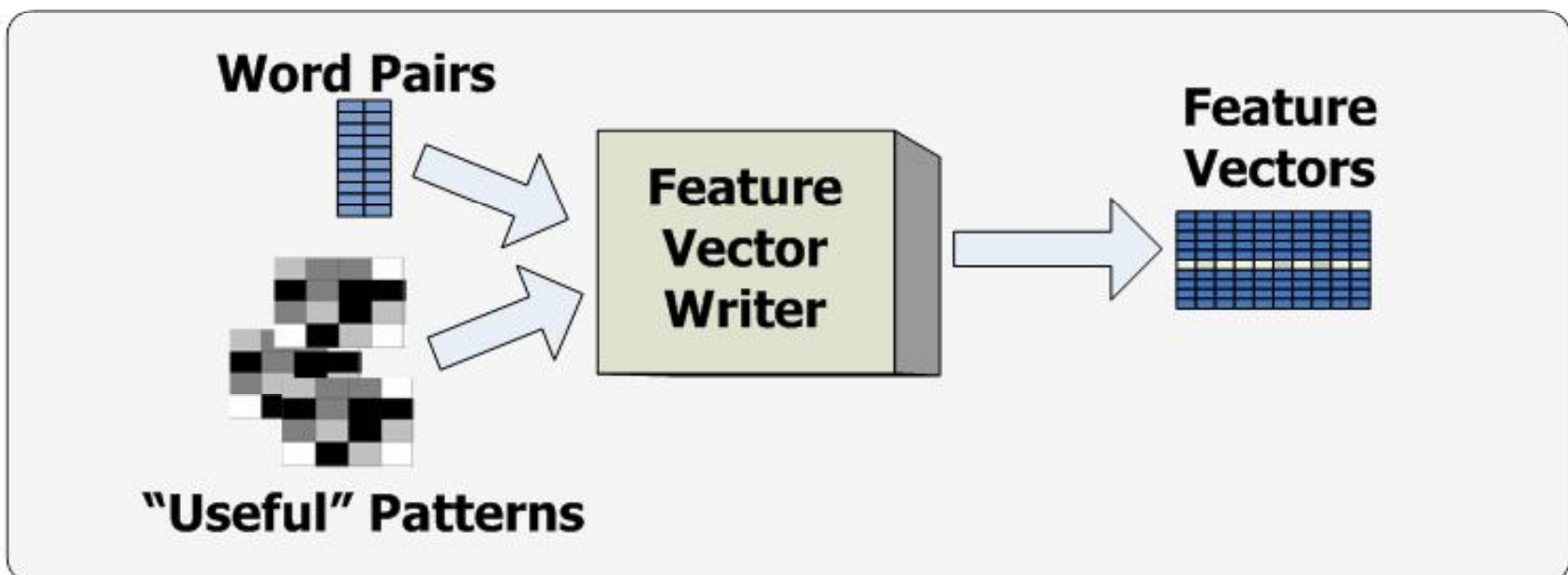
TreeProcessor

- Generates all patterns in a corpus containing word pairs from a given pair set
- Applies a set of patterns to a corpus to generate word pairs
- Most CPU intensive section. Threaded to make use of a user-defined number of CPUs



FeatureVectorWriter

- Determines the features used (The useful patterns based on pattern frequency in the corpus.)
- Generates the feature vectors for a set of word pairs such as the set generated by TreeProcessor



Extending WordNet using Generalized Automated Relationship Induction

Todd Sullivan
todd.sullivan@cs.stanford.edu

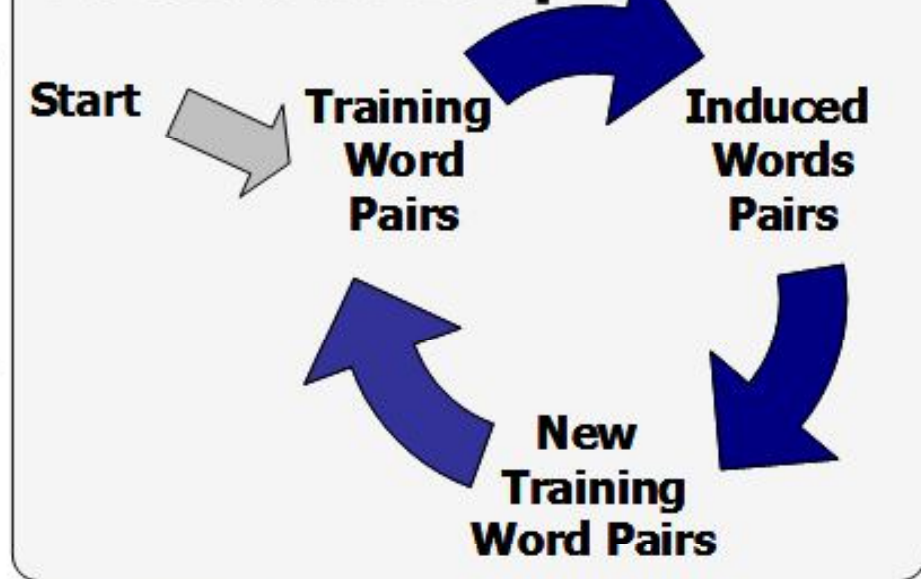
Nuwan I. Senaratna
nuwans@cs.stanford.edu

Lawrence McAfee
lcmcafee@stanford.edu

Classifier

- Logistic Regression
- Naive Bayes
- Support Vector Machine
- Entropy-Based Classifier
- Create your own

Feedback Loop



I love books by Mark Twain and other similar authors.

