UTEP Prosody-Based Property-Inference Package (ProsProp) API

Version 1, June 2017

Nigel G. Ward, Jason C. Carlson, Ivan Gris, University of Texas at El Paso

with support from Darpa’s Lorelei Program, and testing by Anindita Nath

# Purpose

Given an audio file as input, this package returns a set of property-value pairs that characterize it, for example "urgent: 1.7", "medical-assistance: 0.2, locally-relevant 1.3". This is intended initially to support two uses:

1. The LoreHLT evaluation to be run in August 2017. In this case the output properties will be the 11 “situation frame types”. These outputs can be scored directly, or combined with estimates obtained with other methods by other teams.
2. Stance inference, as described in the paper Inferring Stance from Prosody, available at the stance website: [http://www.cs.utep.edu/nigel/ stance/](http://www.cs.utep.edu/nigel/%20stance/) .

The ultimate goal, in either case, is to support downstream, user-facing applications. Usage scenarios include: using these outputs in an interface which enables users to filter or sort news segments by some property, using these outputs to populate the slots of situation frames; using these outputs to support an information-visualization-style interface, where users could view aggregated information across many audio files, seeing for example how the values of certain property vary over time, or with source, or with location, and so on; using these outputs as features input to a classifier to make other decisions, in combination with other features, relating for example to priority.

# Internals

Currently this works by comparing the prosody in the input file, computed over 6-second patches, with the prosody in the annotated reference data, using a k Nearest Neighbors, as described in “Inferring Stance in News Broadcasts from Prosodic-Feature Configurations”, Nigel G. Ward, Jason C. Carlson, Olac Fuentes. Computer Speech and Language, submitted. In future we may use other techniques, using for example neural networks, hopefully without needing to change the API.

# USAGE

Inputs:

1. A pathname. The directory containing audio file to process, either directly or within subdirectories. Each must be in .au format; we’ve tested using mono, 16 bit, 8000 Hz.
2. The name of a prosody-property mappings (ppm) file (.mat format). Each such file encodes the prosody-to-property mappings derived from a set of annotated reference data. These are a shallow encoding, and as such these files are very large. UTEP will produce new ppm files as needed, for example for a new Lorelei incident language, and release them at the stance website. The ppm files can be downloaded from the stance website. To use, just download the file and copy it to the ppmfiles folder. The ppm files currently available are:

* eng-stance-lnews.mat (English, stance, local news)
* cmn-stance-kazn.mat (Mandarin, stance, KAZN local news)
* tur-stance-voacri.mat (Turkish, stance, VOA and CRI Bolt data, LDC2014E115)
* uig-stance-vnews.mat (Uyghur, stance, Lorelei data: various news) \*\*
* spn-stype-vnews.mat (Spanish, situation types, Lorelei data: various news) \*\*
* rus-stype-vnews.mat (Russian, situation types, Lorelei data: various news) \*\*

1. Flags. This is a string containing one or more of the following letters:

* j -- write a json file
* 4 -- downsample by 4 to run 4 times faster
* u -- output user-facing stances instead of raw stances

Output

A matrix of property values, with one row per input audio file and one column for each property. In addition, a json-format file will be written to the outputs directory if the ‘j’ flag is specified. This file is in the format used by the USC evaluation scripts.

Since this is a general-purpose package, the set of properties can be anything that has been previously annotated. Thus, depending on the ppm file used, the package can return values for different sets of properties. The range of output values depends on the property set, and will have the same range used by the annotators of the reference data; thus downstream applications may need to rescale or discretize these for their own purposes. Currently raw stance properties are in the range 0-2 and situation-type properties in the range 0-1.

The code can be involved in various ways:

* In Matlab, by directly calling the prosprop function.
* From Python, by calling prosprop.py, which is a wrapper for the Matlab code
* Through Swagger, once we get it working. This will also enable use as a web service

Sample calls for Matlab and Python are given below.

# Installation

First download the midlevel toolkit from <https://www.github.com/nigelgward/midlevel/> . Then download the PPM package from www.cs.utep.edu/nigel/stance/ppm.tgz \*\*and unpack it.

Test procedure 1:

* Run Matlab
* addpath src
* prosprop(<aufilespec>, <ppmfilename>)

Test procedure 2:

* Install Matlab/Python Engine: <https://www.mathworks.com/help/matlab/matlab_external/install-the-matlab-engine-for-python.html>
* Add the environment variable to the path PythonPath:

C:\Python34\Lib;C:\Python34\DLLs;C:\Python34\Lib\lib-tk; Path: C:\Python34

<http://stackoverflow.com/questions/3701646/how-to-add-to-the-pythonpath-in-windows-7>

* From a console, go to the stance directory and use the following command:

python prosprop audio.au ppmfile

# Properties

The various property sets and their meanings are defined elsewhere; for convenience they are summarized here in the table.

(The 14 raw stances are described in Inferring Stance from Prosody and in Preliminaries to a Study of Stance in News Broadcasts. The 8 user-facing stance aspects described on page 16 of the Stance Usability Packet. All are available from http://www.cs. utep.edu/nigel/stance/ . The situation types are defined in Lorelie Situation Fram Annotatoin Guildelines for Speech Data, v 2.6, by Appen.)

|  |  |  |
| --- | --- | --- |
| **Raw Stances** | **User-Facing Stances** | **Situation Types** |
| 1. Bad 2. Good 3. Deplorable 4. Praiseworthy 5. Controversial 6. Factual Information 7. Subjective Information 8. Unusual or Surprising 9. Typical or Unsurprising 10. Local 11. Prompting Immediate Action 12. Background 13. New Information 14. Relevant to a Large Group | 1. Bad … Good 2. Deplorable …. Praiseworthy 3. Subjective … Factual 4. Typical … Unusual 5. Distant … Local 6. Just Talk … Urgent 7. Old Information … New Information 8. Idiosyncratic … Relevant to a Large Group | 1. Civil unrest or widespread crime 2. Elections and politics 3. Evacuation 4. Food supply 5. Urgent rescue 6. Utilities, energy, or sanitation 7. Infrastructure 8. Medical assistance 9. Shelter 10. Terrorism or extreme violence 11. Water supply |

In future, properties might include

* Status Variables, as described in NIST LoReHLT 2017 Evaluation Plan,which will probably map to 5 scalar properties: past-need, current-need, future-need, relief-already-sufficient, and urgency.
* Emotion Properties, as described in the SEC Pilot Plans document by Kathleen McKeown. These will probably map to 4 scalar properties: positive, negative, anger, fear.

# Validation

You can use the following commands with the test data included in the package to validate the output and ensure everything is properly installed:

**Python:**

python prosprop.py 21d.au eng-stance-lnews

Output: [[0.873445592636224 0.662971158487514 0.198231685803193 0.0918977810705203 0.159899973619268 1.83175798412842 0.592660987129277 0.202377661884140 0.0833049361963311 1.46956478302109 0.619844008398634 1.25386686338000 0.887652596310017 1.21696422038363]]

**Matlab:**

out = prosprop('C:\path\audio\21d.au','C:\path\datasets\eng-stance-lnews.mat.mat');

Output: [[0.873445592636224 0.662971158487514 0.198231685803193 0.0918977810705203 0.159899973619268 1.83175798412842 0.592660987129277 0.202377661884140 0.0833049361963311 1.46956478302109 0.619844008398634 1.25386686338000 0.887652596310017 1.21696422038363]]

In this case, the training data was annotated with stance values. Each value represents a stance in the following order:

If the data was annotated differently, for example, by situation type, running the following code in Python or Matlab respectively

python prosprop.py 21d.au spn-stype-vnews

out = prosprop('C:\path\audio\21d.au','C:\path\datasets\spn-stype-vnews.mat');

Will return a set of 11 type values in the following order:

# SYSTEM REQUIREMENTS

This system has been tested with Matlab R2016a. It should also work with Matlab R2014a or greater.

The python wrapper was tested on Python 3.4; it should also work with Python 3.2 or 3.3, but no higher than 3.4.

# Directory Structure

doc

Documentation, including this file

src

Source code, in matlab, including

-prosprop.m top-level function to infer properties from prosody

-prosprop.py a wrapper to call the above from python

-makePPM.mat, a function to create a prosody-property mapping file

-mono4.fss, a feature-set-specification file

-README.txt, more information on the functions present here

ppmfiles

Models of the prosody-properties mapping, for various languages, data sets, and annotations

aufiles

input audio files to be classified

swagger

# Swagger server and .yaml file for the API provided for reference.

# You may want to regenerate these files using the .yaml to ensure

# proper paths and server generation.

outputs

where the mat file (and sometimes also a json file) for outputs are written

extensions

extra code for experiments etc. outside the main workflow.

# A subset of the information contained in this document on how to run the system

C:\..\stance\README.txt \*\*

# Creating PPM Files

In this API the process of converting annotated training data into a ppm file is not exposed. If you wish to use the system with a new language or training data set, you can either ask UTEP to do it, or you can do it yourself in Matlab, using the provided functions. In overview, you will need to annotate your new dataset, write a parser for your annotation format, compute the prosodic features for the data, and modify the code to associate a descriptor with the new dataset. For this you will require …. (midlevel\*\*?). The functions are individually documented.

A ppm file is just a mat file with the following variables:

* **provenance**: a string describing the source of the mappings. This will specify at least the language, the annotation type (stance or situation types), the feature set, and the date.
* **propertyNames**: a cell array of strings
* **featureSpec**. The name of a feature-set specification (.fss) file, such as mono4.fss. This specifies what features to compute for the input file; these will be the same as those used to create the ppm file. In addition to what is specified, two temporal features are always appended.
* **means**. A vector of means for each of the features, computed over the reference set. This is uses to normalize the input-file features in the same way.
* **stddevs**. A vector of standard deviations, similarly.
* **algorithm**. Specifies which algorithm to use; currently only kNN is available.
* **model**. A model based on the training data, suitable for use by ‘algorithm’. For kNN, the model is of the form ‘segmentData’, which is an array of structs, one per segment. Each struct includes
  + **features** the prosodic-features for each patch of this segment. The rows are patches (one every 100ms throughout the data) and the columns features, as described by featurespec.
  + **properties** a vector of property values, for example 14 stance values
  + **broadcastName** for debug: the audio file or directory this segment belongs to
  + **startTime** for debug: start time in seconds of this segment within broadcast
  + **endTime** for debug: end time in seconds of this segment within broadcast

# Matlab-Python Integration Notes

How to call Matlab functions from Python: [https://www.mathworks.com/help/matlab/matlab\_external/call-matlab-functions-from-python.ht](https://www.mathworks.com/help/matlab/matlab_external/call-matlab-functions-from-python.html)

How to call user scripts on Matlab from Python: <https://www.mathworks.com/help/matlab/matlab_external/call-user-script-and-function-from-python.html>

Some additional configuration may be required to access Matlab from Python in your server. You can find the steps at <https://www.mathworks.com/help/matlab/matlab_external/install-the-matlab-engine-for-python.html>

# History

This is a rewrite of the February 2017 release. The major changes are:

* a simple interface, with a single top-level function to call
* better documentation and comments
* sample data and outputs (see comments in the code)
* support for handling LDC-Appen style corpora, with individual audio files for each segment
* support for both stance and situation-frame types
* addition of temporal features, which improve performance