**Instructions for Robodoc.py**

This script reads in .asc data files that are created by EyeTrack, and outputs .da1 files that are used by EyeDry. It replaces EyeDoctor. It deletes trials with blinks or track losses in your critical region, based on criteria that you specify, as well as trials with problematic display changes.

**Some important things to know before starting**

* Robodoc.py is written in python 3. You will need to have python 3 installed on your computer (and if you also have python 2, you’ll have to be sure you invoke the right python version). Other than that, it is platform independent.
* Python 3 is picky about UTF-8 encoding. If you get an error message that says something about UTF-8, it means there's some unrecognized character(s) in an input file. Try converting your input file to UTF-8. This is easy in TextWrangler and other text editors.
* The program has functionality for dealing with multi-line displays, and for dealing with display changes. However, it does not yet have the ability to do *both* of these things at once. If you have a multi-line experiment with a display change, the program may not work well.
* Multi-line texts will work only if your .asc files have first been run through fix\_align.R, Andrew Cohen's script, which is available on the UMass Eyetracking Lab website.

**What you need to do before running the program**

* You have to put several things in the same directory as Robodoc.py (or, in the path, if you specify a path name):
  + Your parameter file. An example is included with the distribution of Robodoc, named parameters.txt. You may modify this file using any text editor (don't use Word), and you may keep the name, or use a different name. Details of the parameter file are below.
  + The file that provides region boundaries for your sentences (traditionally, referred to as a .cnt or 'count' file). You make this file with make\_cnt.py.
  + If you have display change trials, the experiment script itself.
  + The .asc files that you want to process.
  + You'll need to specify various things within the parameter file. Here they are, in order (more detail is found in the comments in the parameter file itself):
  + Provide the list of .asc files you will be analyzing in the file\_list variable. If you don’t provide a list of .asc files you want to analyze, the program will use all the .asc files in the current directory (the folder where Robodoc.py is stored). In this case, the variable must still be present, specified as an empty array, as in file\_list = []. A list of files that were processed will be written to the files\_processed.lst file.
  + If you are using IDLE or another interpreter that doesn’t allow arguments to be specified, you can tell RoboDoc to print out information about excluded trials by adding this verbose variable to the parameters file: args.verbose = 1.
  + If your experiment has display changes, set DC = 1. Also set the tolerance for excluding bad display changes (i.e., how many ms late the reported display change can be), and provide the name of your script file. (As noted above, display change filtering is not yet implemented for multi-line displays.)
  + Set multi\_line to "y" or "n" depending on whether you have multiline texts. If you set it to "y" you must set first\_line\_y and line\_sep\_y to the appropriate values.
  + Specify lowest\_cond and highest\_cond to the condition numbers you are going to analyze.
  + If desired, set values for excluding trials based on the overall number of blinks, or based on the maximum duration of any blink.
  + If you want to exclude trials based on a blink or track loss in a critical region, set blink\_reg\_exclude to "y". If you do this, set blink\_region to the number of the critical region in your region boundary file, and provide the name of your region boundary file. Note: there can be only one critical region. If you have multiple regions you want to examine, you probably should make one critical region that includes both for the purposes of RoboDoc.
  + Set blink\_gopast to 0 or 1 as appropriate, depending on whether you want to exclude trials when there was a blink during first pass on your region (0) or during go-past (1).
  + Set saccade\_dur\_crit if you want to eliminate trials when no blink was reported but there was a long saccade (e.g. > 100 ms) into or out of the critical region.
  + Set short\_crit to some number other than 1 (e.g., 80) if you want to combine fixations of duration less than that number of ms with the preceding or following fixation, if those fixations are on adjacent characters.
  + Set the auto\_exclude variable to 1 in order to eliminate files that exceed the maximum number of excluded trials (either in total or for each condition) due to blinks. Otherwise, set this variable to 0. The file with the list of files to be excluded will be called exclude.lst. Files that were not excluded will be written to keep.lst.
    - If the auto\_exclude variable is set to 1, you must specify what the exclusion threshold is for each individual condition. For example, if exclude\_threshold = 4, any condition that has a trial with 4 or more blinks in the critical region will be added to a list of filenames to exclude. You can ignore this option by setting the variable to a large number, e.g., 1000.
    - You must also specify a value for the abs\_exclude\_threshold variable, which determines the cutoff point for the number of total trials that can be excluded for each subject. For example, if abs\_exclude\_threshold = 10, the .da1 file will be added to exclude.lst if there are 10 or more total excluded trials, regardless of which condition they were from. You can ignore this option by setting the variable to a large number, e.g., 1000.
  + Set the auto\_exclude\_DC variable to 1 in order to exclude files that exceed the maximum number of excluded trials due to display change errors. If auto\_exclude is also set to 1, then the exclude\_threshold and abs\_exclude\_threshold variables will be used for display change thresholds as well.

**How to run the script**

* There are two ways to run Robodoc.py. The first does not take any arguments at the command line other than the script name itself. For example, say you are running your script from the Mac Terminal. You simply type the following (or whatever invokes python3 on your computer, e.g., python RoboDoc.py, or c:\python3\python RoboDoc.py):

python3 Robodoc.py

In this case, you will be asked to provide the parameter file, at a command prompt.

* The second method takes the parameter file name as a command line argument. You would type the following:

python3 Robodoc.py parameters.txt

* In addition, you can add an optional “verbose” flag (--verbose or -v) to the command in order print out information about the number of files excluded.

python3 Robodoc.py –v

python3 Robodoc.py parameters.txt –v

* Note: if you are using IDLE, you cannot add arguments while running. To use the verbose flag with IDLE, set the args.verbose variable to 1 in your parameters file, i.e., args.verbose = 1.

**What Robodoc outputs**

* A .da1 file for every .asc file, written to a separate da1\_files folder. These are just like the .da1 files that you'd get out of EyeDoctor, except for a few things:
  + If you do not have display change, number of fixations is in column 6 of the .da1 file, and the data start in column 7, in contrast to the previous 8 and 9, respectively. You'll need these values for EyeDry.
  + The total time on each trial now starts when the gaze box disappears (i.e., when the sentence itself appears) rather than when the gaze box appears.
  + Related to this last point, the first fixation on each trial will generally not be a whole fixation: it will be whatever portion of a fixation follows triggering of the gaze box. (E.g., if the gaze box is triggered 150 ms into a 250 ms fixation, the first fixation in the .da1 file will have duration 100 ms.)
  + X and Y values that fall outside the range of the sentence have value -1.
* A .bli file for every .asc file, which provides, for every trial, the start and end time of any blinks. It also reports whether the trial was rejected based on the criteria in the parameter file, identifying rejected trials as having been rejected because of blinks or display change problems. These files are written to a separate blink\_files folder in the current directory.
* A .dc file for every .asc file (again in its own subdirectory), with information about trials lost because of display change problems.
* Several summary files:
  + A file called sum.sub, which provides, for each .asc file, the number of trials that were deleted.
  + A file called files\_processed.lst, which is a list of da1 files that were created. This file can be used in EyeDry.
  + A file called exclude.lst, which is a list of da1 files that exceeded the threshold for number of trials excluded.
  + A file called keep.lst, which is a list of da1 files that did not exceed the threshold for number of trials excluded. This file can also be used in EyeDry.
  + A Blinks-summary.csv file, which lists how many trials were excluded by condition on the basis of blinks. If auto\_exclude = 1 in the parameters file, it will also indicate whether a subject was excluded for having too many trials excluded. In this case, all such files will be written to exclude.lst.
  + A Display-change-errors-summary.csv file, which lists how many trials were excluded by condition on the basis of display change errors. If auto\_exclude\_DC = 1 in the parameters file, it will also indicate whether a subject was excluded for having too many trials excluded. In this case, all such files will also be written to exclude.lst.