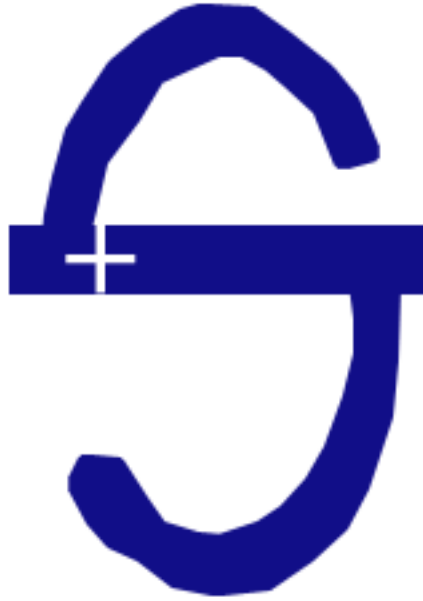


SOKATH Manual

SOKATH version 1.3
SSAUL version 0.5 BETA



Karl G. D. Bailey
December 2003

SOKATH MANUAL

Contents

1. What can Sokath do?
2. Starting Sokath
 - 2.1. Formatting Region Files
 - 2.2. Command line arguments
3. The Sokath Interface
 - 3.1. Menus
 - 3.1.1. File
 - 3.1.2. Data
 - 3.1.3. Analysis
 - 3.1.4. Help
 - 3.2. Regions and Variables
 - 3.3. Input box
 - 3.4. Display boxes
4. Entering data into Sokath
 - 4.1. Selecting a subject
 - 4.1.1. New data file
 - 4.1.2. Existing data file
 - 4.2. Starting a new trial
 - 4.3. Entering a new data point
 - 4.4. Correcting mistakes
 - 4.5. Saving the data file
5. Analyzing data in Sokath
 - 5.1. Trigger files
 - 5.1.1. Format of a trigger file
 - 5.1.2. Loading a trigger file
 - 5.1.3. Setting a trigger
 - 5.2. Setting a time window for analysis
 - 5.3. Setting a condition for analysis
 - 5.4. Selecting an analysis
 - 5.5. Copying data to another program
6. APPENDIX A - Data File Format
7. APPENDIX B - Carving up the visual world
8. APPENDIX C - Analyzing composite scene-fixation videos from the head mounted eyetracker
9. APPENDIX D - The SOKATH String Analysis User Library (SSAUL)

1. What can Sokath do?

Sokath is a program that can be used to score and analyze serial categorical data. This program was developed specifically for the scoring and analysis of data from psycholinguistic experiments using head mounted eye trackers. However, this program can be used for analysis of any kind of data where there are changes in category over time.

This program uses commands from Microsoft Excel and saves data files in Excel format, so it requires Excel to be installed. Sokath has only been tested on Windows XP Pro.

2. Starting Sokath

Sokath can be started by either creating a shortcut, or by using the command line. In either case, the following syntax is necessary:

```
sokath.exe -f regionfile
```

A region file is required to set up the possible categories of response, as well as the experimental conditions and responses to be graphed in the data analysis preview box.

2.1. Formatting Region Files

The region file has the following structure. All text in bold are required by the program.

```
# variable 1 - description  
conditions:ABBR  
# variable 2 - description  
conditions:ABBR  
# variable 3 - description  
conditions:ABBR  
# region definitions  
category:ABBR:a  
# graph  
ABBR  
# end
```

Sokath expects all of these definitions to be present. If your experiment does not have three variables, you will still need to include a single condition in the variables that you are not using.

The abbreviations that are stored in the region file will be stored in the data file, not the long condition name.

The last argument for the region definition is the shortcut key that can be used. Shortcut keys for the variable conditions are assigned by the program.

2.2. Command line arguments

The only command line argument at this time is for the inclusion of a region file.

`-f regionfile`

All region files must have the extension `.rgn`. Replace `regionfile` with the file name, without the extension.

3. The Sokath Interface

The interface consists of four major areas, in addition to a line of response text at the bottom of the window.

3.1. Menus

Menu commands are separated into four sections: those dealing with open and saving data files, those dealing with changes to data files, those dealing with analysis of data, and those that provide help with the program.

3.1.1. File

New Data File - Create a new data file.

Open Data File - Open a Sokath data file. Data files have the extension `.skh`. Either this command or *New Data File* must be the first command of a new session.

Save Data File - Save a data file. This command functions as *Save As* in other programs. Deal with it.

Exit - Exit Sokath. This command will always prompt you to save. Again, deal with it.

3.1.2. Data

New Sub - Create a new subject in the current data file.

Open Sub - Switch to a different subject within the data file.

Delete Last Entry - Deletes the last data point entered, or (more correctly) the last data point listed in the data display box.

Delete Any Entry - Deletes a data point from the current data display box by line number.

New Trial - Changes the number of the current block and trial, in addition to the initial frame number (this assumes that you are using data from a video).

3.1.3. Analysis

Load Trigger File - Load a trigger file (see 5.1).

Set Trigger - Set a trigger for analysis from the trigger file.

Set Time Window - Set the number of frames for inclusion in the current analysis. Sokath currently assumes a frame rate of 30 fps. This setting will be user defined in future versions.

Set Condition - Set the condition for each variable for analysis and display in the analysis preview box.

Momentary (Fix) - Conduct a momentary probability analysis of fixations for the current condition, starting at the current trigger and continuing for the length of the current time window. Raw data is written to the output dialog.

Cumulative (Fix) - Conduct a cumulative probability analysis for the current condition, starting at the current trigger and continuing for the length of the current time window. Raw data is written to the output dialog.

Momentary (Saccade) - Conduct a momentary probability analysis of fixations for the current condition, starting at the current trigger and continuing for the length of the current time window. Raw data is written to the output dialog.

Cumulative (Saccade) - Conduct a cumulative probability analysis for the current condition, starting at the current trigger and continuing for the length of the current time window. Raw data is written to the output dialog.

Make Window Strings - Output all states that occur during a particular time window. These strings can be input to SSAUL for further analysis.

Make Region Strings - Output all states that occur during a particular trigger region. This can be used with SSAUL to conduct the type of analysis proposed in the Appendix of Altmann & Kamide (in press).

Output - Open output dialog.

Toggle Updating - Toggle whether display boxes are updated. Due to a known issue with redrawing the display boxes, this can speed up analyses.

3.1.4. Help

About - Display information about Sokath.

3.2. Regions and Variables

The region list can be used after a frame value is entered into the input box in order to enter a data point into the data file.

The variable lists are used to set the current condition for data points.

For all lists, commands can be entered by clicking on the radio button, or by using the shortcut key.

3.3.Input box

Frame numbers are entered in the input box. The Esc key returns focus to the input box.

3.4.Display boxes

The data display box displays the raw data for the current subject. This can be used for checking data, or determining which line number to enter if a line needs to be deleted. When an analysis is run, the current subject will cycle through each subject, ending with the last subject in the data file. This is a feature, not a bug, and can be turned off using the Toggle Updating command in the Analysis menu.

The analysis preview box displays a preview graph when an analysis is run (except for the Make Strings commands). The color of any graph line corresponds to the colors in the region list. Only those regions listed in the region file # `graph` list will be graphed.

4. Entering data into Sokath

4.1.Selecting a subject

Before data can be entered, a subject must exist.

4.1.1. New data file

Creating a new data file automatically creates a new subject.

4.1.2. Existing data file

Data can either be entered for the current subject, or a new subject can be added to the data file.

4.2.Starting a new trial

The trial definition should be changed to the current block and trial numbers as well as the initial frame. Variable lists should also be updated.

4.3.Entering a new data point

Enter the current frame number in the input box, then either click on the appropriate region, or press the appropriate shortcut key.

4.4. Correcting mistakes

If an incorrect data point is entered, it can either be immediately deleted by deleting the last line, or be deleted later by deleting any line.

4.5. Saving the data file

Remember to save! Sokath does not save a temporary backup file.

5. Analyzing and outputting data with Sokath

5.1. Trigger files

5.1.1. Format of a trigger file

A trigger file is an Excel spreadsheet that has a minimum of five columns. The first column contains trial numbers. The next three columns contain information that associates lines with conditions ('x' is a wildcard). The fifth (and any other) columns contain trigger times in seconds.

See the sample trigger file included with the program for more information.

5.1.2. Loading a trigger file

This involves loading a file. The program does all of the hard work for you.

5.1.3. Setting a trigger

This is even easier.

5.2. Setting a time window for analysis

Pick a number of frames. Currently, thirty frames equal one second.

5.3. Setting a condition for analysis

I trust that you can figure out how to do this. Make strings command output all conditions simultaneously.

5.4. Selecting an analysis

Choose either a momentary or a cumulative probability graph. The momentary probability graph gives the probability that a certain region is currently being selected (e.g. fixated, pointed to, touched) at a given time after a trigger. The cumulative probability graph gives the probability that a certain region has been selected by a certain point after a trigger.

5.5. Copying data to another program

Right clicking in the output dialog box will allow you to select all and copy the data into another program for graphing. This is especially important for the Make Strings commands.

6. APPENDIX A - Data File Format

A new sheet is created for each subject. Each sheet contains eight columns. They contain the following information (in order) for each data point: subject number, block number, trial number, frame number, region abbreviation, variable 1 condition, variable 2 condition, variable 3 condition.

7. APPENDIX B - Carving up the visual world

The trickiest part of analyzing head mounted eyetracking data is determining how to carve the visual world. This is simpler in psycholinguistic experiments, given the rather sparse nature of the 'scenes' involved. However, experiments involving vones may be more difficult to carve. The following papers give examples of how this is accomplished.

[papers go here]

8. APPENDIX C - Analyzing composite scene-fixation videos from the head mounted eyetracker

This appendix assumes that a video clip exists for each trial.

- Determine the frame that matches the initial (zero) trigger.
- If necessary, change the current condition
- Change the trial definition (CTRL-t).
- Step through the video
 - Whenever there is a change of region being fixated, enter a new data point.
 - A fixation in a region is defined as two consecutive frames with the cross hairs in approximately the same area, taking into account drift due to the eye tracker moving, eye moving,

- microsaccades, etc. Use fixations on the central dot to determine offset.
 - If there is a clear change of direction, but the crosshair only rest for a single frame, this may still be scored as a fixation.
 - Fixations on a boundary between two regions should take into account direction of eye movement.
 - Fixations in non-regions are scored as OTHER or UNKNOWN
 - If a blink occurs during a fixation, it is not scored.
- The trial ends when an appropriate move is completed, and an eye movement is made, or arbitrarily when a new trial is about to begin.

9. APPENDIX D - The SOKATH Sequence Analysis User Library (SSAUL 0.5 BETA)

SSAUL is a command line program that takes strings from SOKATH as input and outputs tables of data to a file for further analysis in a statistic program. Currently, SSAUL must be in the same directory as the input file. Output files will be placed in the same directory.

The prompts from SSAUL are as follows:

Input File? (This is the input file containing the strings output by SOKATH. Delete any trailing newlines.)

Region Code? (All region codes of interest should be listed, separated by the pipe "|". For the "Simplify String" option, this line should contain all regions to be included in the simplified string. This line is ignored by the comparison analyses.)

Output File? (This is the output file where the data table will be saved.)

[1] Subject or [2] item analyses? (This determines how the data will be averaged for analysis. Enter "1" or "2" as necessary. This is ignored for "Simplify Strings", and the "Pre Cluster" analyses.)

The program then lists available analyses in the library. It then prompts for the user to select an analysis.

Select an analysis? (Enter the line number of the analysis that you wish to run.)

For the single string comparison analyses, the following prompt will also appear.

Pattern to match? (Enter the pattern with each code separated by a tab.)