Liaison CS Meeting Agenda – Thursday 9/19/24

1. Update on work from last week–we have been working primarily on the statement of work which is under Prof. Hope’s review now and will be submitted to liaisons for initial review before Tuesday.
2. Questions
   1. Is there any unannotated waveform data (question by Milo)
3. Description of mosquito data intake process from Zach (5 mins)
4. Items from Dr. Backus
5. Finish teaching students about aphid waveforms and WinDaq (WWB) –
   1. Compressed recordings show problems in the last 1/3 of most files – decide what to do
   2. Early difficulties with labeling and Backus corrections
   3. Show “descriptions of files” document and explain
   4. Go through examples and exemplars I made
6. How to correct some remaining errors
   1. P-E1’s were all removed, reducing errors listed by half. Remaining errors are:
      1. Repeats (e.g., P-P) – should they be deleted or added? Students decide
      2. Sequential errors – usually mean an event in between was missed/not measured – what to do?
   2. Can we figure out a non-manual way to correct remaining errors? Or not worth trying?
7. Demonstrate labeling one or two aphid WWB files
8. Reserve 5-10 mins for contributions from Dr. Reif and Dr. Cooper
9. Reserve 10 mins for final questions from the Mudd team

Minutes

* Update from Mehrezat on work and progress last week
  + Statement of Work will be sent tomorrow, Thursday, or Monday to Liaisons
    - Dr. Backus requests for us to send as Word file
* Is there unannotated data?
  + Zach clarifies that WinDaq files that are “just lying around” could be useful.
  + Dr. Backus confirms that this exists
  + Dr. Cooper states existence of similar data and we will let her know if we want it
* Zach describes his process of converting WinDaq files
  + Used a script to convert/combine WinDaq files with label file
  + Clarification that this was done using label files provided by Dr. Cooper and Dr. Reif
  + Assurance that it’s easy for us to work with WinDaq files and manipulate them to be useful for us
* Demonstration of Windaq & explanation of data errors by Dr. Backus
  + Shown waveforms were recorded 20 years ago; there are many problems (recorder was new to the process, sets are differently labeled, etc.)
  + Baseline is very noisy, but centered around 0 (done manually by Dr. Backus before our meeting)
  + Terminology
    - Stylets: mouth parts inserted into the plant
    - Probing: stylets move in and out of the plant in search of food
  + Very important to recognize probing spikes against baseline noise
  + Showed waveform labeled “D” instead of “C” due to keystroke issues
  + P-D (potential drop?) alternation goes on for a great portion of the beginning of the recording
  + E family waveforms
    - “potential drops” = large negative spikes marking the beginning of E behaviors
    - E1 = run of quick positive spikes (moving around cell wall)
    - E2 = run of quick negative spikes (penetrating cell wall)
  + E, F and G are seldom, but go on for a long time
  + Showed Excel file describing the first waveform file in Experiment 1 (has been shared with us)
  + Dr. Backus points out general downward decline of waveform center (quality of recording declines)
    - Dr. Cooper describes occasional adjustments of hardware variables during recording
    - DC loss will make it more difficult to identify waveforms in the latter end of the data
  + Glue could come off and interrupt a recording
  + So do we want to discard data sections?
  + October 4 2004 file is incorrectly labeled (sections over or under measured)
    - But Dr. Backus relabeled the file correctly
  + “4-keystroke” annotation
  + TBF values are critical to know how accurate the annotation is
  + Very difficult to differentiate between ET and E2 sections
  + Zach identifies negative duration value (always mistakes, must be corrected)
  + Exercise: in descriptions file, look at Channel 3 description (exemplar recording) outcome and compare to other channels
  + Errors file: had (has?) a lot of false errors, removed by Dr. Backus
    - Tracks human errors by detecting event transitions that don’t make sense
* Dr. Cooper
  + It’s exciting that there is so much aphid data, but because there are so many issues with the data would it be more useful to begin with the mosquito data?
    - Zach: mosquito data is already in a state where it is usable; aphid data is going to take much longer to put in a usable form
    - Dr. Backus: says she emphasized the worst case scenarios in the interest of time, most of the files are more accurate (especially in the later ones)
  + Would our program be able to mark inconsistently labeled data?
    - Zach: probably not, since it’s designed for initial annotation
* Zach: data shape seems to be more focused on, are frequencies also important?
  + Dr. Cooper: it would be really cool if the program could show the frequency. Frequency is a very important aspect of the waveforms.