CTrax analysis notes

Track analysis workflow:

Track input

Select and upload CTrax files into R objects

to do: define a `trx` class for the data frame and some wrapper class for the experiment list

Script 0.1:

- input dir and file information therein
- report file list and name list
- output `raw` list

Select and binarise the light protocol from a csv file

to do: automate threshold

Script 0.2:

- input light protocol file and protocol structure info
- report constructed time protocol (vs raw data)
- output `t` time protocol data frame

Data processing 1 - global

Garnish data combining time protocol information

to do: impose correct var type (factor, int, num) to variables

to maybe do: add some vars (log transforms and abs), think of more?

Script 1.1:

- input `raw` and `t`
- report: not sure, as this is a behemoth of data. Could do the track integrity and a general track plot by id~genotype
- output `garn` list

Data processing 2 - quadrant protocol

Calculate PI (by fly and by repeat)

Script 2.1:

- input `garn`
- report boxplots of PIs by genotype
- output `PIfly` and `PIrept` (`PIbyfly` and `PIbyrept` too?)

Data processing 3 - section3 protocol

Tidy section 3 time protocol

generalise to a given known protocol (frames on, frames repeat)

add drift?

Script 3.1: *(note: there is no reason not to run this immediately after 0)*

- input `t` and protocol parameters
- report maybe time offsets and such
- output `tsynth` data frame and perhaps the `progress` time annotator

Extract section 3 to produce a subsetted data frame of all genotypes

to do: once the exp class is defined, create a listtodf method

to do: think more seriously about the structure of `tsynth` (esp the strange chunk 3 declaration)

Script 3.2:

- input `raw` or `garn` and `tsynth`
- report something like mean trends of various variables over the protocol? or a summary of data (which vars, how many ids per genotype). or the non-binned data could be generated here (as in var paths over time, overplotted by id and blink), or that could be a separate report or script.
- output `s3df`

Collapse stats to 1s bins

to do: look at median and standard deviation too

Script 3.3

- input `s3df`
- report boxplots of meaningful variables by second and repeat
- output `s3mean` (and possibly `s3median` and `s2df`)

Figure design

try to work out what this is good for, if anything

show plot of PCA for a laugh

Script 3.4

- input `s3df`
- report path patterns by second and repeat
- output `s3figs`

Data processing 4: Virtual quadrant design

Generate synthetic fly paths by fragmenting and collating *on* and *off* real paths

idea: define x and y in radial coords from the centre of the plate (somewhere in `garn` there may still be an `arena` list, or I can pull out its definition) to simplify track assembly by rotating things about instead of shifting

to do: decide/try either chunking by classifier or going entirely blindly.

Data processing 5: Classifier

Describe fly behaviour by its vars

pick up old classifier again

try unsupervised, with new vars (spin, possibly friends)

look into markov chain analysis of behavioural transitions