Overview of Valection

Chris Cooper & Dorota Sendorek

February 5, 2018

Contents

1 Overview

1.1 Description

Valection contains a variety of algorithms for choosing verification candidates from competing tools or parameterizations, to fairly assess their performance against each other.

This software requires the valection package (http://labs.oicr.on.ca/boutros-lab/software/valection).

1.2 Synopsis

There are six selection methods available through six functions. They all take the following arguments:

- **budget**: an integer specifying how many candidates to select

- **infile**: a path to a file which contains the calls from all callers

- **outfile**: a path to a filename where the calls should be outputted

- **seed** (optional): an integer to seed the random number generator with (used to randomize sampling)

1.2.1 Formatting the infile

The infile should be formatted with a tab separating the caller and call on each line:
    caller1 name\ta call this caller made
    caller2 name\ta call this caller made

e.g.
    magnifying glass   chr1 576834
    magnifying glass   chr1 6878924
    eye dropper   chr1 496267
    eye dropper   chr1 6878924

Note that the call can contain a tab, but the caller may not.

1.3 Functions

The functions are named as follows:

- run.directed.sampling

- run.random.sampling

- run.equal.per.caller

- run.equal.per.overlap

- run.increasing.with.overlap

- run.decreasing.with.overlap

1.4 Usage

```
> require('valection');
> # run the sampling to select 10 candidates
> run.equal.per.caller(
+       budget = 10,
+       infile = "/home/me/calls.valec",
+       outfile = "/home/me/selections.txt",
+       seed = 50
+       );
```

1.5   Acknowledgements

- Paul Boutros, PhD, PI - Boutros Lab

- The Ontario Institute for Cancer Research

1.6   Copyright

This software is copyright (c) 2015 by the Ontario Institute for Cancer Research.