NAME

ssampler − sample random flux distributions from the stoichiometry of a metabolic network.

SYNOPSIS

ssampler [*options*] [−o <FILE>]

DESCRIPTION

Ssampler (stoichiometry sampler) is a tool for sampling uniformly distributed, random flux distributions from the stoichiometry of a metabolic reaction network. The specification of the reaction network is assumed to be contained in a FluxML document. The generated flux distributions are saved to a HDF5 file. **Ssampler** is used to generate initial feasible flux distributions and random samples for distributed multi−start optimization.

COMMON OPTIONS

−h, −−help

Show a brief help for all command line options.

−i, −−in <FILE> [default: stdin]

The name of the FluxML (XML) input file. If omitted, the FluxML document is expected on standard input.

−o, −−out <FILE>

The name of the HDF5 output file. This option is required (HDF5 files are never written to standard output).

−c, −−configure <CFG> [default: 'default']

Because FluxML documents may contain several **<configuration/>** elements this option allows to specify the configuration that should be used for the simulation. If this option is omitted it is assumed that the FluxML document contains a configuration with the name "default".

−L, −−list

Specifying this option results in a list of allowed configuration names for the specified FluxML document. The program exits immediately after emitting the list.

−l, −−log DEST

Specify the destination for the internal logging. In the most simple case **DEST** is a file name of a log file. In case the file exists new log messages are appended. Apart from log files it is possible to publish log messages to file descriptors, UNIX domain sockets, UDP and SCTP ports, and a small graphical user interface.

A file descriptor is specified by **fd:[num]**, where **[num]** is the number of the file descriptor.

A unix domain socket in the local file system is specified by **unix:[name]**, where **[name]** is the name of the socket file

A UDP or (connectionless) SCTP port is specified by **[proto]:[host]:[port]**, where **[proto]** is either "udp" or "sctp" and **[host]** is the name of the destination host and **[port]** is a UDP or SCTP port number on the destination host. Please note that the length of log messages is bounded by the minimum safe UDP packet size − log messages containing more than 548 characters will be truncated.

Finally, log messages can also be sent to a small GUI by specifying the destination **@gui@**. The GUI requires a working Perl/Tk installation and a running X server.

In order to capture all log messages concerning the command line processing this option should be specified in front of all other options.

−v, −−verbose 0..10 [default: 5]

Specify the verbosity 0, 1, ..., 10 of generated / emitted log messages. The meaning of the different log levels is as follows:

- **0 (QUIET)** do not emit log messages at all.
- **1 (ERROR)** only emit severe error messages.
- **2 (WARNING)** only report severe errors and warnings.
- **3 (NOTICE)** report all errors and warnings including important informal messages.
- **4 (INFO)** report all errors, warnings and all informal messages.
- **5 (THROW)** in case of an exception, try to give a diagnosis of the error; sometimes even gives a backtrace of the current function stack.
- **6 (DEBUG0)** emit the more important debugging messages.
- **7 (DEBUG1)** emit the less important debugging messages
- **8 (DEBUG2)** emit the superfluous debugging messages
- **9 (DEBUG3)** emit annoying debugging messages.
- **10 (DEBUG4)** don't dare to use it!

SAMPLING OPTIONS

−n, −−nsamples <NUM> [default: 1000]
The number of samples to be generated. Defaults to 1000 if omitted.

−s, −−sample <n,x,nx> [default: n]
This option is used to specify what to sample. Possible values are 'n', 'x', and 'nx' − for sampling only net fluxes, only exchange fluxes, or both net and exchange fluxes. The default is to sample only the distribution of net fluxes.

−S, −−sampler <g,h> [default: g]
Use this parameter to specify the sampling method. The default sampling method is Gibbs sampling ('g') which usually gives the best results. Alternatively, Hit−And−Run sampling ('h') may be used (slightly faster).

−b, −−bound−net <VALUE> [default: 100]
The bound for the magnitude of unbounded net fluxes. Defaults to 100.

−p, −−bound−xch <VALUE> [default: 100]
The bound for unbounded exchange fluxes. Defaults to 100.

−f, −−faster
Especially if the metabolic network is very large, the sampling process may be very slow. In this case the user has to option to skip the warmup phase performed after each step of the sampler. This results in much faster sampling. However, the generated samples may be of lower quality, i.e. less uniformly distributed.

−r, −−randomize <NUM> [default: 0]
This option allows to specify a randomization interval. After the specified number of steps the sampler is re−initialized with a randomized start point. Specify a value greater than 0 if you have the feeling that the generated random samples are not uniformly distributed, which is sometimes the case for a degenerate flux space. This option defaults to 0, i.e. randomization is disabled.

EXAMPLES

Sample a random net flux distribution consisting of 1000 random samples using the Gibbs sampling method (most simple case):

ssampler −i network.fml −o samples.h5

Sample net and exchange fluxes and generate 10 flux distributions. Bound the net fluxes to +/−15.5 and use the Hit−And−Run sampling method:

ssampler −i network.fml −o samples.h5 −n 10 −p 15.5 −S h −o network_new.fml

SEE ALSO

fwdsim(1), fitfluxes(1), sscanner(1)

AUTHOR

Michael Weitzel

This manpage was written by Michael Weitzel <mich@el−weitzel.de>.