Variable Neightbor Search (VNS) documentation

VNS is a program permitting to find epistasy patterns using Variable Neighbor Search.

Compilation instructions

In order to compile this method a makefile was given at the root of VNS's folder. Please notice that a general makefile was generated at the project's root permitting to produce executables of both methods.

Preparation of pre requist

In the makefile please change BOOST_FOLDER value with path of the installed BOOST library on current workstation.
Moreover, please check if current g++ version is compatible with C++11 functionalities (version >= 5.0). Please also check BOOST library. Authors cannot guaranty compatibility with other BOOST version even if BOOST version seems to be compatible. For this method version 1.61.0 was employed.

Compilation

To compile this method, please call those lines at the root of the project:

```
cd vns
make
```

I a recompilation is needed use:

```
cd vns
make clean
make
```

Execution of program with:

```
./vns.exe <path_to_genotype_dataset> <path_to_phenotype_dataset> <path_to_parameter
```

Summary of files of this project

Module	Files	Usability	Inputs	
Parsing	file_parsing	Allows to parse file containing data set	Files names given as argument in program's call	
Parsing	parameters_parsing	Allows to parse parameters and save them in parameters class	Will use ./parameters/parameters.txt	
VNS	VNS	Core of this project allows to run VNS's algorithm	Parameters object containing all parameters Data set object containing: - Matrix of genotype data - Vector of phenotype data	
			Matrix column of genotype	

Statistics	statistics	Used to do g2 test of independence	(subset of genotype matrix with only tested SNPs) Vector of phenotype Indexes of tested SNPs	
Statistics	contingencies	Build observed and expected contingencies table, some functions useful to work on contingencies tables	Matrix column of genotype (subset of genotype matrix with only tested SNPs) Vector of phenotype Indexes of tested SNPs Number of observation in tested subset	
Output writing	output_writing	Can save algorithm's results in file	Patterns and associated score	

Parameters

Parameters of this program are stored in parameters folder into a file called parameters.txt. This file is commented to permit an easy setup for particular needs. This allows to tweaks many parameters like output directory or output prefix and many parameters specific to VNS.

Launch the analysis of the toy data set

With this project is provided a python script allowing a simple launch on every platform equipped with python. This script will launch program on a naive data set with default configuration given by authors.
To execute this script please call:

```
./launch_vns_toy_example.py
```

Expected output

This method give as output a txt file gathering all results.

```
# Result from vns
# Pattern || occurences || chi2-score || p-value || unreliable case
{M0P7,M0P8} || 4 || 87.8235 || 2e-16 || 0
# Execution time : 116 milliseconds
```

Those results can be checked using our evaluation tool given in evaluation folder.
Using a toy data set:

```
Filename,TP,FP,FN,recall,precision,f_measure,power
genotypes_toy_dataset.txt,2,55,43,0.04444444444444446,0.03508771929824561,0.961754
```
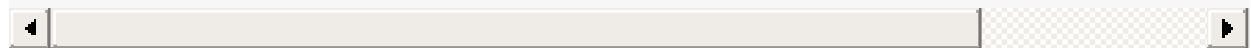
Built With

- BOOST - peer-reviewed portable C++ source libraries

Authors

Tanguy Lallemand -
Jonathan Cruard