# Timothy (short) User Guide

Maciej Cytowski (ICM, University of Warsaw)

January 8, 2016

## 1 Introduction

This document describes Timothy user details. Timothy is a novel, large scale computational approach which allows for cell colonies simulations to be carried out over spatial scales up to 1cm in size (more than $10^9$ cells) i.e. the tissue scale.

Please note that Timothy was designed as a model generator. Basic simulations of cellular colonies can be executed by adjusting the parameter file which structure is described in this document. More specific simulations (e.g. special tumour growth scenarios) sometimes require minor source code modifications. If you are interested in running more advanced simulations please contact the code's author: `m.cytowski@icm.edu.pl`.

## 2 Prerequisites and compilation

There are few prerequisites needed for the successful installation of Timothy:

- C compiler supporting OpenMP (tested compilers: GNU gcc, IBM xlc),
- MPI library (tested implementations: OpenMPI, IBM POE for Power7 AIX, IBM MPI library for Blue Gene/Q),
- Zoltan v.3.8 library (available at `http://www.cs.sandia.gov/zoltan/`),
- Hypre v.2.9.0b library (available at `http://acts.nersc.gov/hypre/`),
- SPRNG v.2.0 library (available at `http://www.sprng.org/`).

Installation and configure flags used for installation of the prerequisites:

1. Zoltan example configure command (Intel compilers):
   ```
   ./configure CC=mpiicc CXX=mpiicpc FC=mpiifort --with-id-type=ulong --enable-mpi
   --prefix=/home/user/zoltan/3.81
   ```
2. Hypre example configure command (Intel compilers):
   ```
   ./configure CC=mpiicc CXX=mpiicpc F77=mpiifort --enable-bigint --with-MPI --with-openmp
   --prefix=/home/user/hypre/2.10
   ```

## 3 Parameter file

Standard usage of the program is achieved by defining necessary simulation parameters in the parameter file. An example parameter file is presented in Figure 1. Here, we describe the meaning of all important parameters, allowing better understanding of the functionality of the simulator.

```
# Parameter file for Timothy

SCSIM 0
BVSIM 0
BNSIM 0

NC 1
NSTEPS 128
#524288
DIM 3
MITRAND 1
SECPERSTEP 16000
MAXCELLS 10000000

SIZEX 64
SIZEY 64
SIZEZ 64

#RSTFILE step00000600.rst

RSTRESET 1
OUTDIR results

VISOUTSTEP 1
STATOUTSTEP 1
RSTOUTSTEP 200

COUTTYPE POV
FOUTTYPE VNF

NHS 195000000
TGS 1

RNG BM

MAXSPEED 1

# Cell cycle parameters - healthy tissue
G1 11.0
S 8.0
G2 4.0
M 1.0

# Cell cycle parameters - cancer cells
CG1 6.0
CS 8.0
CG2 4.0
CM 1.0

V 0.5
RD 0.1

# Global fields
GFIELDS 1
GFDT 16000
GFH 16.0
GFLOGDIR log

OXYGEN 1
OXYGENDC 1.82e-5
OXYGENBC 0.1575e-6
OXYGENICMEAN 0.1575e-6
OXYGENICVAR 0.0
OXYGENCONS 8.3e-17
OXYGENPROD 8.3e-1
OXYGENLAMBDA 0.0
OXYGENCL1 5.5e-18
OXYGENCL2 1e-18

GLUCOSE 0
GLUCOSEDC 1.82e-5
GLUCOSEBC 0.1575e-6
GLUCOSEICMEAN 0.1575e-6
GLUCOSEICVAR 0.0
GLUCOSECONS 8.3e-17
GLUCOSEPROD 0.0
GLUCOSELAMBDA 0.0
GLUCOSECL1 5.5e-18
GLUCOSECL2 1e-18

HYDROGENION 0
HYDROGENIONDC 1.82e-5
HYDROGENIONBC 0.1575e-6
HYDROGENIONICMEAN 0.1575e-6
HYDROGENIONICVAR 0.0
HYDROGENIONCONS 8.3e-17
HYDROGENIONPROD 0.0
HYDROGENIONLAMBDA 0.0
HYDROGENIONCL1 5.5e-18
HYDROGENIONCL2 1e-18

TEMPERATURE 0
```

Figure 1: Example parameter file for simulation starting from a single cell placed in the center of computational box.

The first group of parameters defines the most important settings of the simulation. The meaning, expected values and important features of these parameters are listed below.

**WARNING:** list below is not up to date, will be fixed soon.

- `NC`: sets the initial number of cells in the simulation,
  - expected value: positive integer number,
  - simulation can start with a positive integer number of cells,
  - cells are disturbed randomly in 3-D space,
- `NSTEPS`: sets the number of iterations in the simulation,
  - expected value: positive integer number,
- `DIM`: sets the dimensionality of the simulation (2-D or 3-D),
  - expected value: 2 or 3,
- `MITRAND`: controls the direction in which daughter cells are shifted from the center of mother cell during mitosis,
  - expected value: 0 or 1,
  - 1 indicates random placement,
  - 0 indicates placement consistent with the direction of movement of the mother cell in last iteration of the simulation;
- `NHS`: sets the number of cells needed to activate apoptosis i.e. programmed cell death,
  - expected value: positive integer number,
  - apoptosis is activated when the number of cells in the simulation is greater than `NHS`,
- `RD`: sets the probability for each cell of being marked for apoptosis,
  - expected value: real value,
  - assumption: $0 \leq \texttt{RD} \leq 1$,
- `TGS`: indicates that the simulation is a tumour growth simulation,
  - expected value: 0 or 1,

The second group of parameters is used to define the size of the computational box. This can be achieved by setting three positive integer numbers (two in the case of 2-D simulations) corresponding to the size of the computational box in X, Y and Z axes, parameters `SIZEX`, `SIZEY` and `SIZEZ` respectively.

The third group of parameters is associated with setting of time steps and the mesh size used for the discretization of equations describing global fields. Below is the explanation of those three parameters.

- `SECPERSTEP`: sets the number of seconds per each iteration step,
  - expected value: positive real number,
  - we recommend that the value of this parameter is not greater than 3600 seconds (1 hour) since the time of cell cycle phases are usually defined in hours,
- `GFDT`: sets the time step $\Delta t$ (in seconds) for the discretization of equations describing global fields,
  - expected value: positive real number,

```
# Restart parameters
RSTFILE step00007000.rst
RSTRESET 0
```

Figure 2: Parameters required to start the simulation from a restart file.

  - important assumptions: `GFDT` $\leq$ `SECPERSTEP` and `SECPERSTEP` $= k \cdot$ `GFDT` where $k$ is a positive integer

- `GFH`: sets the mesh size $h$ for the discretization of equations describing global fields,
  - expected value: positive real number,
  - unit is a multiple of the maximum size of a biological cell, e.g. for $h = 2$ the mesh size will be equal to the maximum size of a biological cell multiplied by 2,

The next group of parameters is used to define cell cycle phases lengths. There is a distinction between healthy and cancer cells. Mean cell cycle phases lengths for healthy cells are set with the use of `G1`, `S`, `G2` and `M` parameters. In the case of cancer cells `CG1`, `CS`, `CG2` and `CM` parameters are used. The expected values are positive real numbers defining the lengths of each of the phases in hours. There is an additional parameter $V$, which is the variance used to define the exact cell cycle phases lengths for each cell individually.

The last group of parameters is used to control how results of the simulation should be stored. There is an assumption that all output files are located in a single directory. The name of this directory is defined with the `OUTDIR` parameter. In order to control the numerical results of global fields computations user can specify the name of the directory in which the Hypre library log files will be placed, this is defined with `GFLOGDIR`. The rest of the parameters specify how often the output informations are written, i.e.: the `VISOUTSTEP` parameter defines how often the visualization files will be written, the `STATOUTSTEP` parameter defines how often the statistics of the simulation will be written to the standard output stream, the `PROFOUTSTEP` parameter defines how often the profiling information (showing timings of different procedures and functions) should be written to the file called `prof.out`. All these parameters are defined by specifying the number of iterations, which separate successive writing out of the results.

The application allows users to use the checkpoint/restart mechanism. In order to write checkpoint files the user has to specify how often this should be done. This is done by setting the parameter `RSTOUTSTEP` to the number of iterations of the simulation, which separate successive checkpoints. In order to restart the simulation from the checkpoint file the parameter file should be modified to include two additional lines presented in Figure 2.

The parameter `RSTFILE` should be set to point to the file containing the restart information. In such a case, by default all simulation parameters are set to values stored in the restart file. However sometimes it is useful to use a starting field containing a large scale model of a healthy tissue to start a completely new simulation with some of the parameters overwritten. This can be achieved by setting the parameter `RSTRESET` to 1.

# 4 Running the simulation

In order to run the simulations it is required to define the number of parallel processes and number of OpenMP threads assigned to each process. The application should be executed with the use of a special command used for running MPI programs. This command is system dependent, e.g. `mpiexec` is usually available on Linux clusters but on the IBM Blue Gene/Q architecture `runjob` or `srun` commands should be used. Here we present an example of execution commands used on the IBM Power 775 system:

```
setenv OMP_NUM_THREADS 2
mpiexec -n 32 ./timothy -p parameters.txt
```

At the beginning of the simulation the application reports the most important parameters by writing the output header to standard output stream. Users should verify that all parameters including the number of processes and threads were set properly.

# 5 Visual analysis of results

Simulation results can be analysed with the use of the visual analysis and visualization package VisNow (available at `http://visnow.icm.edu.pl`). For large scale simulations we have also developed an alternative method based on the ray-tracing visualization package Pov-Ray (available at `http://www.povray.org`). Pov-Ray input files containing 3-D scenes descriptions are created during simulation. Those input files are processed on the HPC system afterwards to produce high quality images (an example script `povray.sh` shows how to run Pov-Ray).