

HealthQ Technologies

Scientific Research – Phase 2B (Pilot Study)

Feasibility Report

Instructions:

- Click on File -> Make a Copy.
- Rename the document to:
`<ProjectName>_Phase2BFeasibilityReport_YYYYMMDD_<Status>`
e.g. AcuteStress_Phase2BFeasibilityReport_20150510_DRAFT
- Click (again) on FILE -> Move to Folder:
CDR -> Implementation -> Research & Development -> Requirements & Specifications
- You can now complete the document. Remove the grey informative text (including these instructions) and replace with your own content.
- Once the draft is completed change the name to
`<ProjectName>_Phase2BFeasibilityReport_YYYYMMDD_FOR SIGNOFF`
- Once the document is signed off by all signatories, upload the document under the name
`<ProjectName>_Phase2BFeasibilityReport_YYYYMMDD_FINAL`

Project Name:	Model Speed-ups		
Project Lead:	Anton de Villiers	Date:	2015/12/08

1. Hypothesis/Research Aim/Research Objectives

The fundamental understanding of the intricacies of the Virtual Human Machine (VHM) is of core importance. The VHM is modeled using Systems Biology which encompasses the computational and mathematical modeling of complex biological systems. In this case the most relevant biological systems within the human body is modelled. The VHM is modelled in Modelica which is an object-oriented, declarative modeling language for component-oriented modeling of complex systems, such as the VHM.

The literature provides many mathematical equations (or systems of equations, often containing differential equations) which may be used to accurately mimic processes within the human body. Such systems of equations are used in the VHM and connected amongst one another to provide a realistic interpretation of the human model.

Upon adding additional components and connections between the underlying systems of the VHM, the execution times related to simulating the VHM naturally increases. In it's current compilation, the VHM simulations execute too slow for practical roll-out to the public domain. It is of great importance to reduce the execution times of the VHM, while still maintaining realistic and accurate simulation results.

A number of ways have been identified to improve the execution times of the VHM. The first area of focus is to make efficient use of the built-in functions of the modeling language of the VHM (Modelica). These built-in functions encompass a wide variety of model building settings, solvers, data manipulators and tolerances that may be adjusted for improvements.

Furthermore, it is important to identify which components within the VHM have a substantial impact on the execution times of the simulations. Once these components have been identified, alternative mathematical equations may be introduced to improve the execution times. It is often computationally expensive to solve Differential Algebraic Equations (DAE) by means of numerical methods (which the Modelica solvers have to perform). Such systems of equations may in some cases be replaced systems of Ordinary Differential Equations (ODE) to reduce the execution times.

A final consideration is to identify whether it would be beneficial to completely remove certain processes within the VHM. Such processes would necessarily have very little impact on the overall impact on the VHM during simulations and should constitute substantial savings in execution times of the VHM.

2. Data Collection Process

Data is collected from the latest VHM model containing the most comprehensive processes within the human body. The data from this model is used as the ideal model and is used as the reference data for any possible new model that has quicker execution times.

Subsequent models that may possess shorter execution times of simulations are compared to this ideal model to identify whether these novel models perform similar to the ideal model. Some threshold information will then be used to identify which changes to the model will be deemed as valid and acceptable to create a new model that would, in turn, be considered to be ideal.

3. Data Analysis Process

Simulation results are compared between the ideal VHM model and altered models which are expected to provide quicker execution times while maintaining high levels of accuracy when compared to the ideal model. A number of critical simulation variables are identified which necessarily have a significant effect on the VHM. The differences for such variables between the ideal model and all altered models are compared for each time step during the simulation of the VHM.

Simulations may be set up for various parameters and initial conditions. To date, all tests have been performed for a specific set of initial conditions whereby the human under consideration starts at resting bodily conditions and the engages in a 100% activity level for 100 seconds.

Prediction error between the models can be measured by the Mean Absolute Deviation (MAD), Root Mean Square error (RMSE) as well as other measures. Another novel measure utilised consists of dividing the two observations at a given time unit, and determining the logarithmic value of that value. All these values are then summed for each time step and the result is divided by the number of time steps considered. Under ideal situations this measure should approximate to 0.

The prediction errors are used to determine which components may cause substantial strain on the execution of the VHM. It is currently assumed that components in the VHM that deviate substantially from the ideal situation when solver tolerances are lowered may be the areas where the model execution is substantially hindered.

4. Results and Assessment

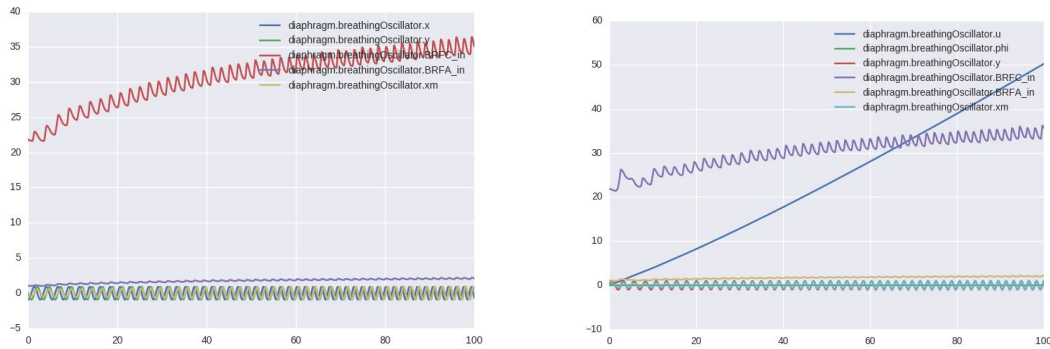
The usage of the built-in functions of the Modelica solvers and compiler has seen significant changes to the execution time of the VHM. In particular, the reduction of the tolerance has brought about a decrease of the execution times by half of the original time required by the ideal model. However, reducing the tolerances does cause some of the variables to deviate from the ideal model. Furthermore, when the tolerances are lowered too much, certain constraints are not adhered to, which provides inaccurate simulation results.

Besides from considering the aspects relating to compiling and executing Modelica code, we can also consider changing mathematical components within the VHM model. To assess the feasibility of the current model investigation an example will be presented for the current Breathing Oscillator in the Diaphragm. The oscillation is determined by two equations as follows:

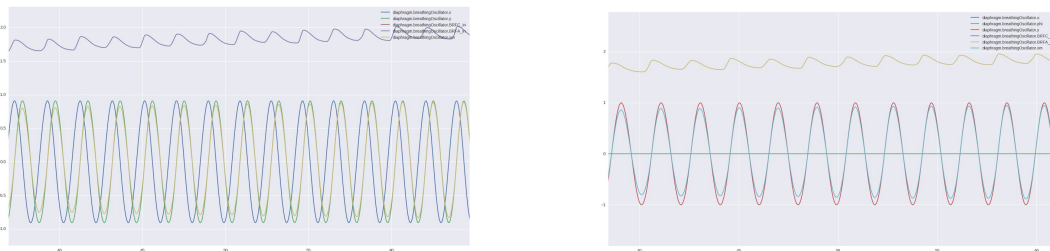
$$\dot{x} = -2 \cdot \pi \cdot f \cdot y$$

$$\dot{y} = 2 \cdot \pi \cdot f \cdot x$$

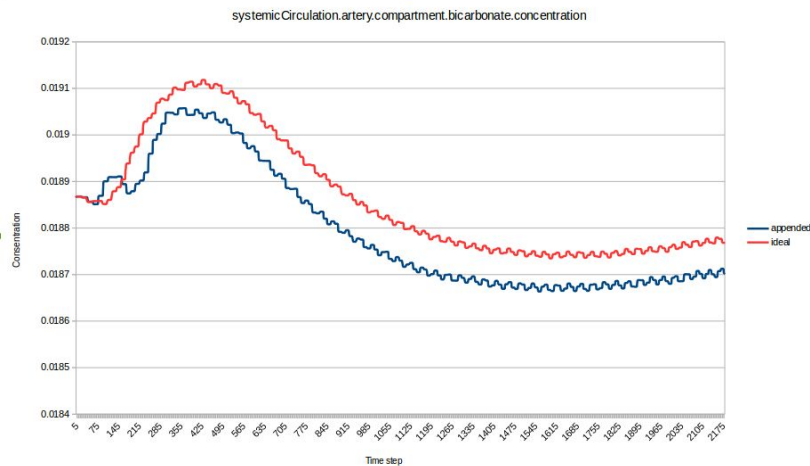
Instead of using the above set of differential equations, it is possible to model these equations by sinus waves that is able to accurately approximate the numerical results of the differential equations. The results of the appended model (containing a sinus wave) can then be compared to the results of the ideal model.



The left image above contains the numerical results of the ideal model, while the image above to the right contains the numerical results of the appended Breathing Oscillator model. By examining the oscillations closely, we are able to make informed decisions on whether this new appended model adheres to the required level of accuracy. The models are run for 100 seconds (horizontal axis).



The images above are present a zoomed viewed of the ideal model (left) and altered model (right) during approximately 35 seconds and 65 seconds into the simulation, but focussing on the oscillations. If such changes are feasible, other components affected are evaluated to ensure that the model still provides the expected behaviour in other components. Below is a graph showing the deviation in the bicarbonate concentration between the two models discussed above. It is important to investigate why this change is occurring and to find a resolution that guarantees accuracy of the appended model.

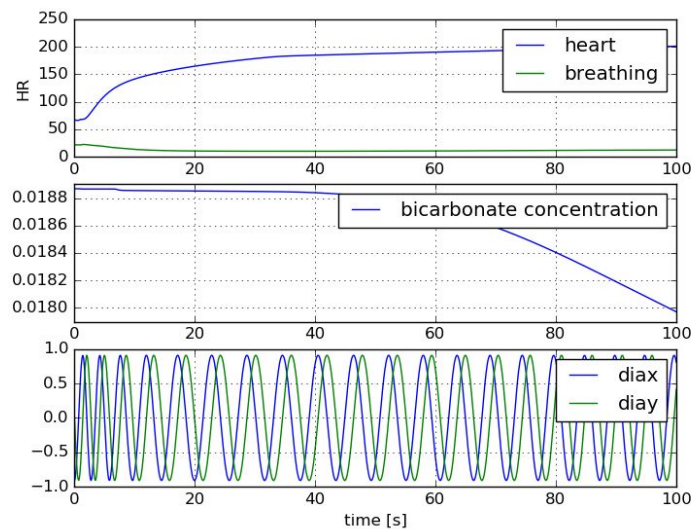


For a standard Intel(R) Core(TM) i7-4710MQ CPU @ 2.50GHz computer containing 8 processors and 16GB RAM, using operating system Linux Ubuntu 14.04, the ideal models simulates 100 seconds of activity in approximately 60 seconds. By modifying standard built-in functions, compiler changes, tolerance changes specific with regards to Modelica and the C implementation of the Modelica model, the time has been reduced to 30 seconds, while maintaining accurate results. The execution time was further improved by incorporating the sinus wave function in the Breathing Oscillator compartment, which has reduced the computation time to approximately 23 seconds.

These improvements have been found after only a single month of investigation with the aim to considerably improve the execution time of the VHM model simulations. At the moment the model seems to be hindered by the numerical calculations of nonlinear equations that must be solved analytically.

During this recent research period there are also been many failed attempts to improve the VHM model. These results, however, should be noted for future reference. The first of these failed attempts at improving the execution speed of the model was creating an Functional Mock-up Unit (FMU) and using Modelica's built-in function to import an FMU (of the VHM) and then simulating the model. The problem with the importing function is that Modelica recreated a .mo file from the provided .fmu input file, defeating the purpose of making use of the FMU to store the data and results in computer memory. In this case no improvements in execution time was reported.

A subsequent approach was to utilise the load_fmu module from the pyFMI library (within the Python programming language). This module will then utilise the FMU and provide results based on simulating the FMU. The current problem is that pyFMI library is only capable of using a single solver to model and solve ODEs, and that solver being CVode. The model executes in a very short period of time, but the results are not correct and do not model the DAEs correctly. Currently the only solver capable of solving the VHM is DASSL. Other solvers either break during simulation times or provide inaccurate results due to the DAE systems not being solved correctly. The pyFMI library can solve the model in approximately 6.5 seconds (for 100 seconds simulation time), but the results are not consistent with expected body activity as can be seen from some of the results displayed below, where the heart rate reached a level in excess of 200 beats per second, the period of the breathing oscillation gradually increases continuously and the bicarbonate concentration experiences a constant decline throughout the simulation.



One other aspect that may bring about improvements to the model is to make use of Dimensional Analysis to determine whether the model may benefit from reducing the number of quantities used. Also, by forcing certain variables to have nominal values close to 1, many numerical rounding errors may be avoided. For example, by multiplying a very large positive number with a very small (decimal) positive number may result in intensive numerical calculations as well as many decimal values that must be rounded. By multiplying two "scaled" numbers who are approximately close to the number 1, the number of decimal values may be reduced and the number of numerical calculations may be simplified.

5. Conclusions and Future Work

By continuing to identify components where the model struggles to perform numerical calculations to solve differential equations and finding simpler or more efficient ways to model such components, savings in execution times should still be possible.

Another consideration would be to perform an investigation into the generated C code upon compilation of the Modelica VHM code. We have very little understanding of how Modelica creates low-level C code which essentially solves the model numerically.

Simulations are currently executed making use of a CPU's computational power. GPUs may provide a means of performing execution of the VHM in much shorter times. A CPU core can execute 8 32-bit instructions per computer time clock, whereas a GPU's can execute in excess of 3200 32-bit instructions per computer time

clock. The cost, implementation and model adaptations will have considered as an initial investigation into determining whether the usage of GPUs will be beneficial.

Jaco Conradie has identified that he has been in contact with some of the Modelica developers, and that they may visit our offices next year. These developers may have more insight into reducing the execution speed of the model, which would be of great importance to HealthQ.

Sign-off		
Proceed to next phase:	YES	NO*
If YES, indicate what the next phase is:	<i>E.g. Phase 3A (Research & Development - Protocol)</i>	
If YES, sign-off meeting for next phase scheduled for:	<i>YYYY/MM/DD</i>	
Designation	Name & Surname	Signature
Project Lead		
Project Coordinator		
Technical Expert <i>(if applicable)</i>		
Project Architect <i>(if applicable)</i>		
Chief Scientific Officer		

***If NO**

Details surrounding the decision.
Normal text: Arial, 9pt, black