A HANDBOOK FOR CovidSIMVL 200 #1507

(“CovidSIM Multiverse Viral Load”)

Ernie Chang, MD PhD Victoria BC

**INTRODUCTION**

**This is the 4th of a series of CovidSIM agent based models written in Javascript for execution on browsers. It is open source and free for anyone to use or adapt, under the general terms of GNU Open Source licensing.**

**All CovidSIM models use agents to represent persons, and the key factors are person size and mobility (which we refer to as “mingle” in this document). The third important factor is density (the population in relation to size for a fixed universe).**

**FACTORS AFFECTING THE COURSE OF AN EPIDEMIC**

**Size is also called Hazard Radius – the larger this is, the more likely that two agents, moving within a confined space (an arena) will end up touching or overlapping, in which situation infection may occur, if one is infective and the other is susceptible. Clearly the smaller the Hazard Radius of an agent, the less likely it is to contact another.**

**The Hazard Radius is therefore the resultant of any number of factors in the context of a contagion-based epidemic – for example, age reduces susceptibility so it will tend to decrease the Hazard Radius. Comorbid conditions such as obesity and diabetes will tend to increase the likelihood of being infected, so will be a positive factor for size. The viral load of a person will tend to create more infectivity, equivalent to an increase in size. The mitigation policy of mask wearing can decrease the susceptibility of an agent, so it is a negative size factor.**

**The algorithms that combine the multiplicity of factors are discussed in a following section.**

**Shelter in place, and isolation, can serve both to limit the mobility of an agent, and can be represented either or both by an effect on size, as well as an effect on movement.**

**Mobility or “mingle factor” is used to propose a move for an agent in a time cycle. The CovidSIM algorithm generates a Pareto-distribution for a move, and then modifies that by the various mingle factors. For example, a person who works as a server in a restaurant or bar moves around more than a client at a dinner table. Thus, the mingle factor for the server would be higher. The modelling of a factor like poor air circulation in an enclosed space can be modelled by giving every person in that space an addition to their individual mingle factors.**

**TEMPORAL DYNAMICS AND VIRAL LOAD**

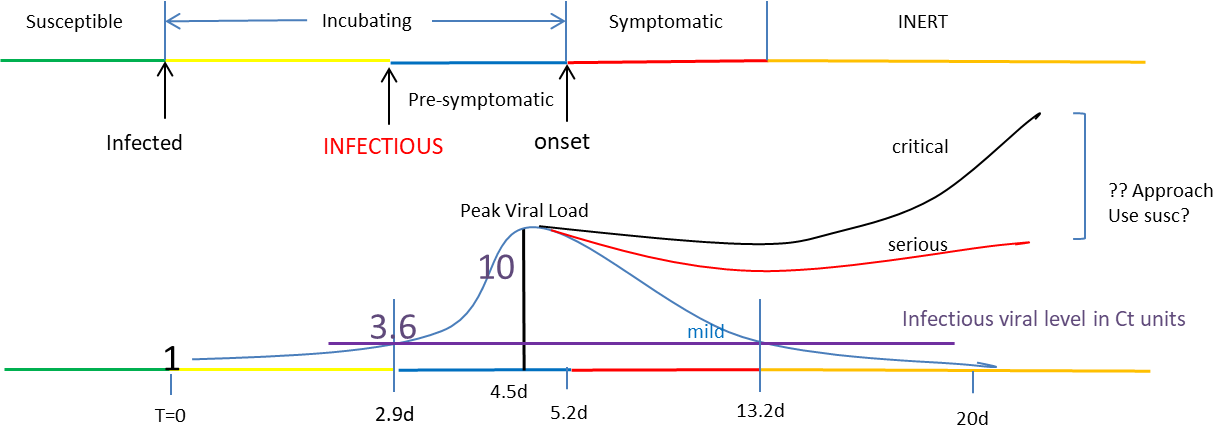
The paper by Xi,He et al. *Nature Medicine* **26,** 672-675(2020) is the basis of a temporal and viral growth model for Covid which is used in CovidSIMVL).

Susceptible

Incubating

Symptomatic

INERT



The time-lines can be changed, and they are applied stochastically for each person, to a +/- randomization of 5%. However, decisions had to be made as to whether the viral load took precedence over the temporal model. The arbitrary decision is to adhere to the time posts, but to also use viral loads as guides to viral growth exponents both from infection to peak, and from peak to undetectable, with 13.2days as loss of infectivity. This approach produced compound growth rate of 1.069 per 0.1days to peak, and a value of 0.865 per day from peak to 20days.

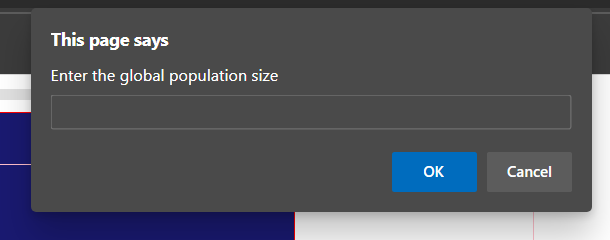
The approach when two agents overlap is to increase the viral load of the agent with a smaller viral load, provided the transfer makes sense in a physiologic way. For example, someone past the infective state (color “orange”) is assumed not to increase their viral load even if they overlap with a heavily infective person.

The viral states of a person progress from “green” – susceptible to “yellow” infected to “blue” – infective to “red” – symptomatic and infective to “orange” – no longer infective and presumed inert (maybe immune or deceased).

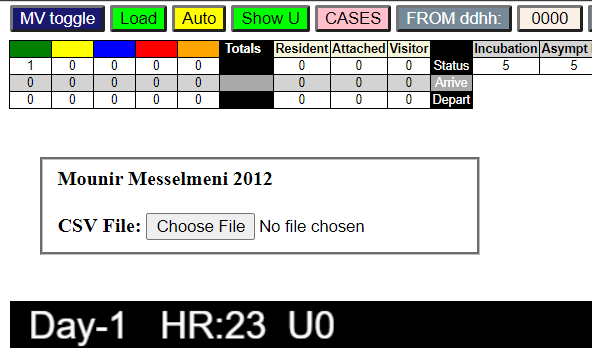
**OPERATING CovidSIMVL**

CovidSIMVL supports nine interacting universes, in the sense that agents can move at specified times from one universe to another, adopting a particular role (one of “resident”, “visitor”, or “attached” – like staff or office worker) with a specific mingle factor.

To start, a total population size covering all the universes is entered, so the system can create universes and skeletal agents.

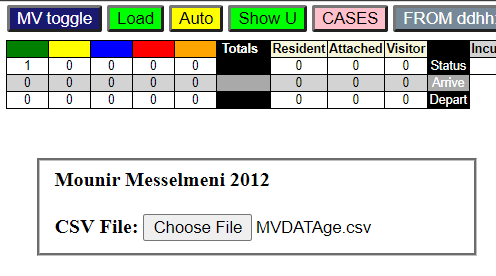


There are two files, at present design…one csv file designates the movement of persons at particular times from one universe to another. Persons are given unique identity numbers (eg starting from 0 to N), and times are in DDHH or just HH format. The Age-Group for that person is also an entry for this first CSV file.

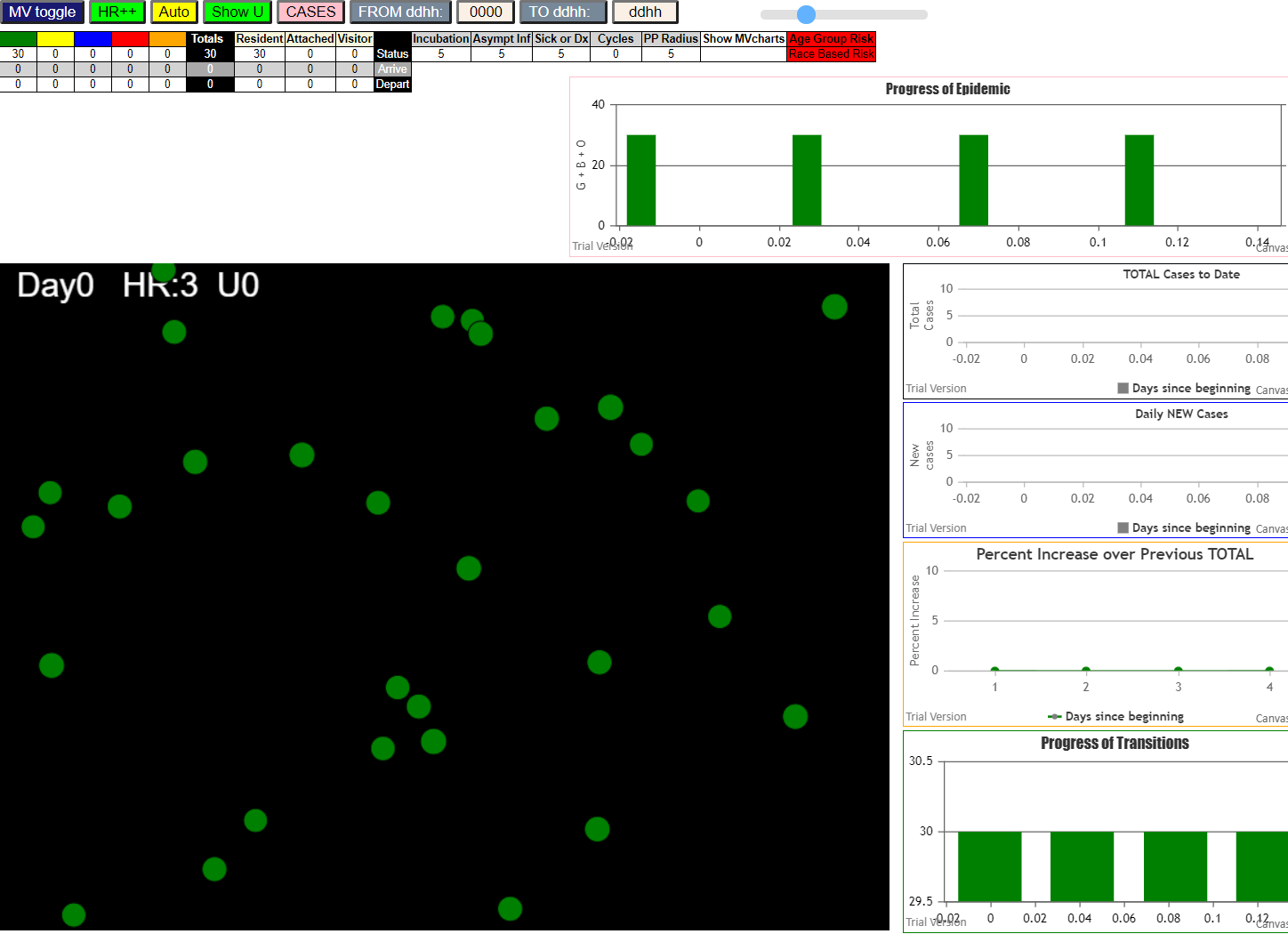


The next csv file is a modifier of selected agents of the population, and assigns a viral status (“blue”), a Time-since-infection to that person, and a viral load from 0 to 10, as well as a single susceptibility factor which is a combination of various personal attributes known to affect the risks of persons with respect to Covidl. These factors include obesity, smoking, chronic diseases, race(?) etc.

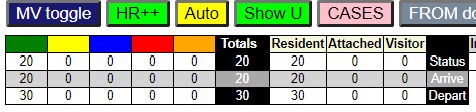
After the first CSV file (“MVDATAge.csv”) has been loaded, the user must select the pink “CASES” button to call up the file open widget that now shows that MVDATAge.csv has been opened, and click on Choose File for the next.



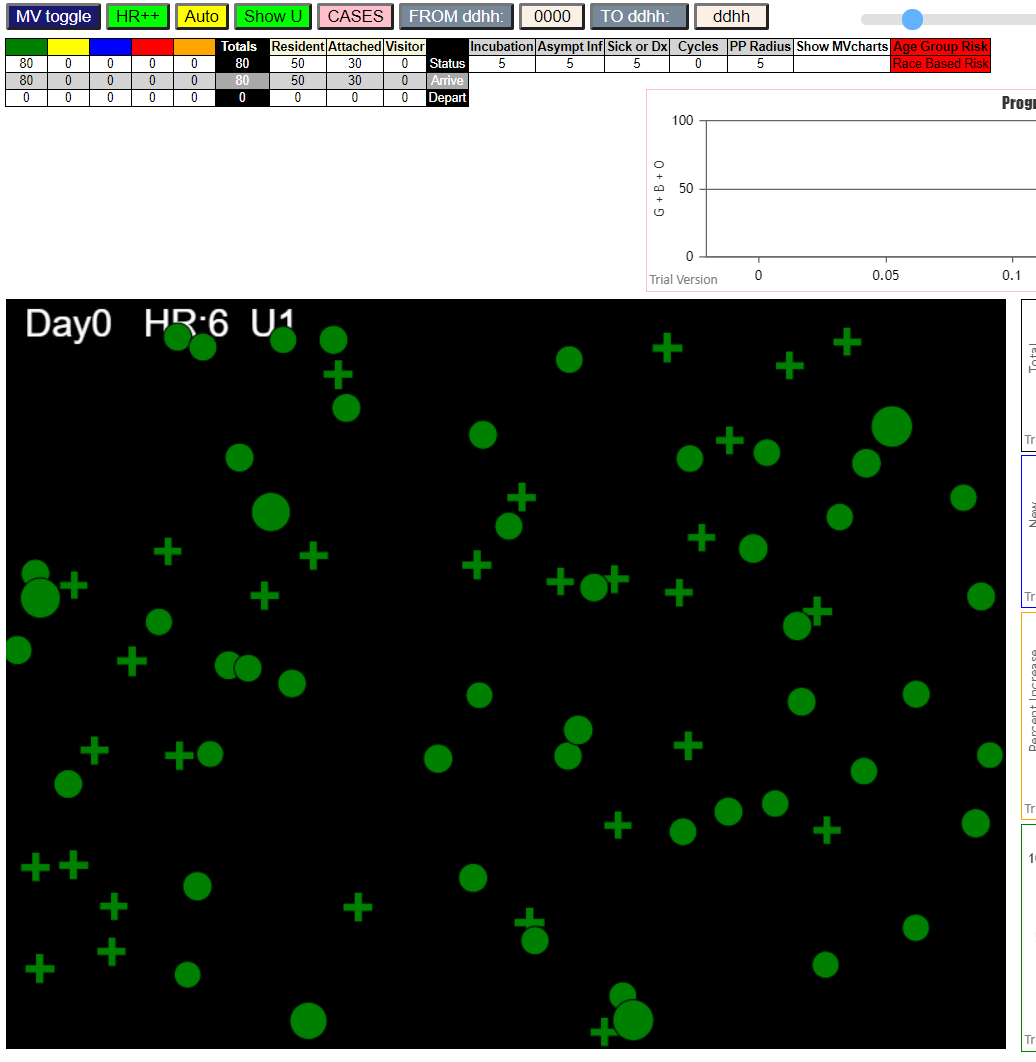
Next, the GREEN “LOAD” button must be clicked, and this will transform to show HR++. One can use this to advance the hours, or use the yellow “AUTO” button to automatically advance the clock. The clock can be stopped by clicking on the AUTO button again. The screen shows time Day and Hour, and the agents in their viral states in the universe shown here as “U0”.



Note here that not only do the agents show up, but the graphs appear, and with each click they will advance accordingly. The smallish table under the menu bar shows the viral states and counts for the universe, as well as the numbers of agents of various classes arriving and departing at the hour, if there is such traffic. For example, we see what happens in hour 6, at which time 30 persons leave for U1 (Long Term Care), where U0 is a universe called “Homes”.

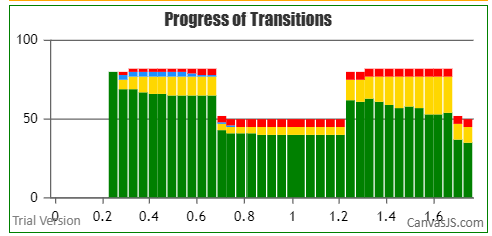


Here we see that Universe U0 has 20 persons, with 30 having left and 20 new arriving. The green “Show U” button will take the viewer to the universe U1 for the same time slot, showing a total population of 80 with 30 new arrivals which are “attached” and 50 “residents” arriving from system initialization.



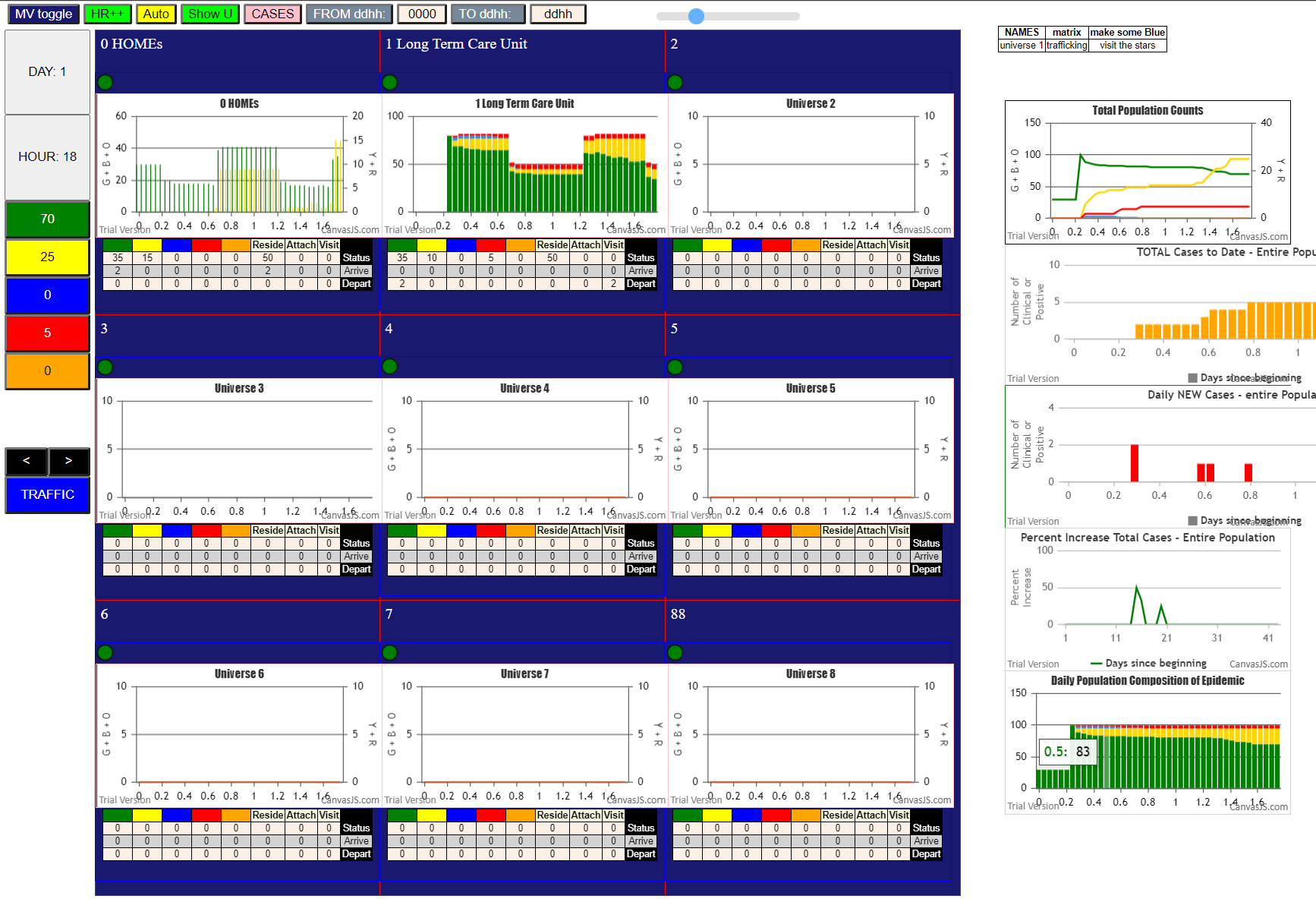
A look at the next hour shows that with viral growth, with the passage of time for the Cases modifications fro the second CSV file, we now have 2 reds (symptomatic), 6 yellows (infected and incubating) three blues (infective), and a number of crosses representing “Attached” roles for staff, while the spheres are permanent residents, and squares will be “visitors”.

By Day 1, Hr 18, we have a situation in the stacked column “Progress of Transitions” which shows



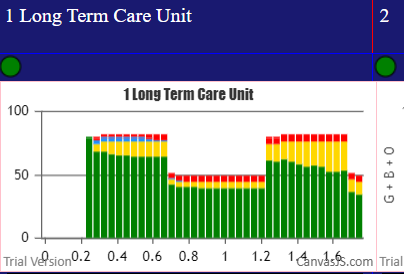
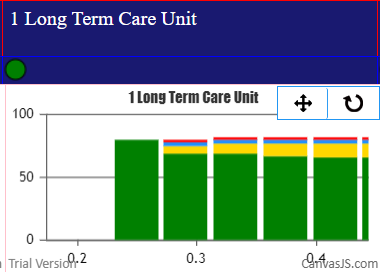
For each hour, the total population in Universe 1 and the numbers of different viral states for the persons in the Long Term Care universe. When the 30 staff leave at 1700, the residents are left on their own (in this case), , and we see in the second day that the new staff of 30 arrivals added significantly to the numbers of incubating persons, which grow in the day through mingling (movement), which takes place as a parameter of five cycles per hour. Then they leave.

The next screen shows the Universes together in one screen, with statistics and a graph for each Universe. Here, we only have the two. The Multiverse iview is invoked by the dark blue “MV toggle” button top left (dblclk).

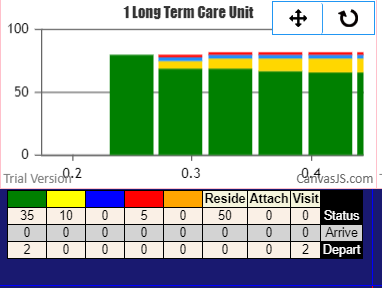


**We see here the side tallies on the left for the entire population, regardless of where each person is, at the present Day-Hour, and the graphs on the right are also with reference to the whole population.**

**For the first two universes (Homes=0 and Long Term Care =1) we have a bar graph for Homes, and a stacked column graph for the Long Term Care. The first graph on the right is a line graph equivalent to SEIR, the next is for Total Cases (symptomatic) in the population, the next for Daily New Cases, the next for % increase day over day for Total Cases, and the last the Stacked Column graph for the entire population.**

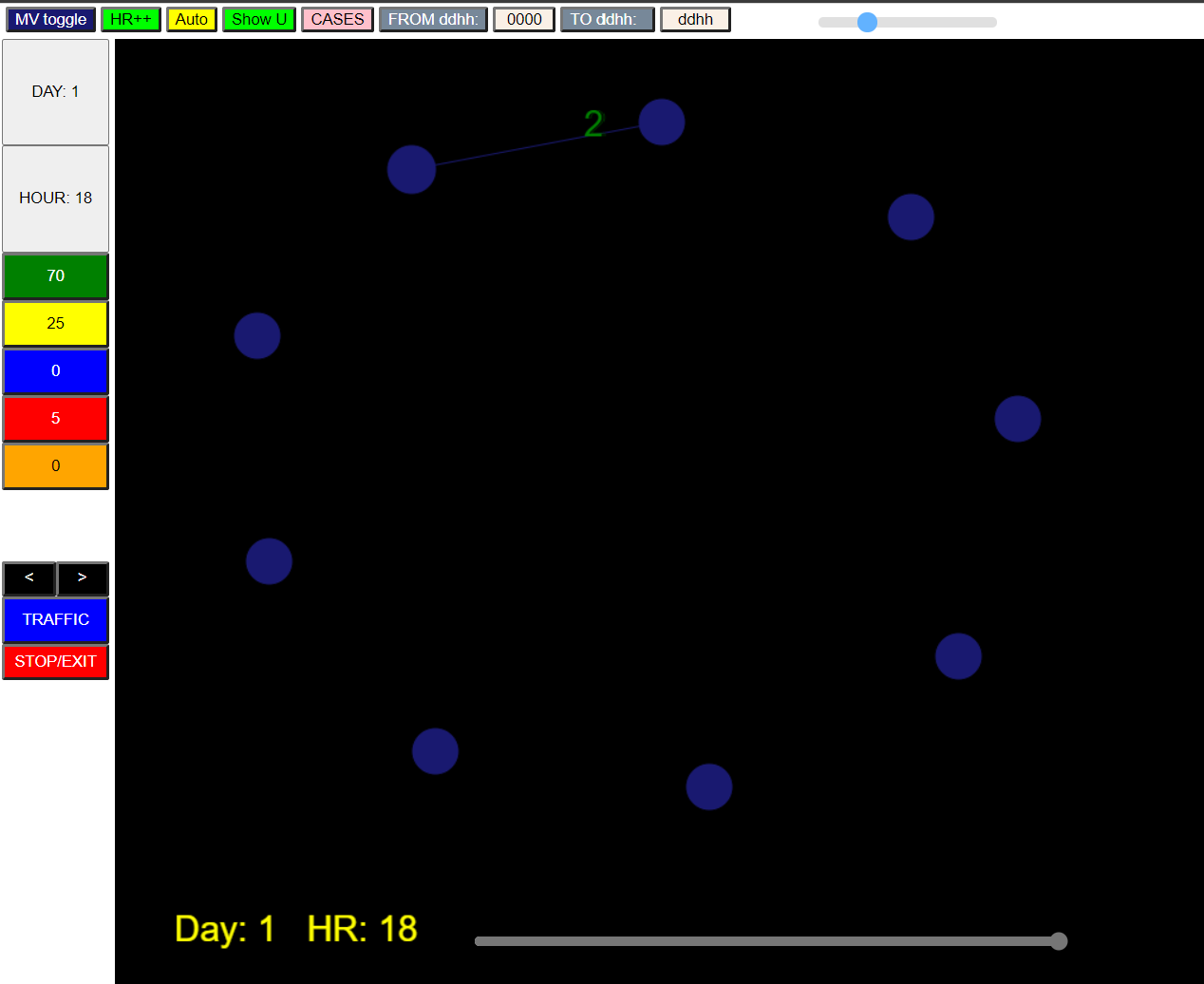
**Each graph can be zoomed. For example, the Long Term Care stacked column can be zoomed to show details of the first few hours on the left, as seen below, by clicking and dragging the mouse over the area.**

**As shown in the single universe view, each Universe has a population and activity tally for the current DayHour.**

****

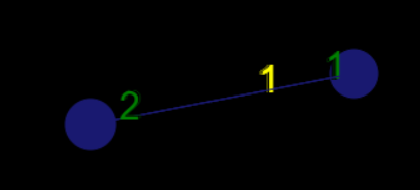
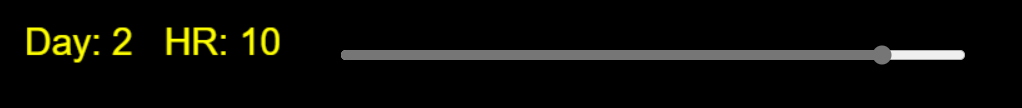
**Thus, Universe 1, at Day1 HR18, has 50 Residents, 35Green, 10Yellow and 5Reds, and has 2Visitors who have departed, and their status was Green.**

**There is another view of the dynamic flow of traffic between the universes, and the intensity and type of the traffic. If one clicks on the “TRAFFIC” blue button, the following screen replaces the previous view:**

****

**The green number “2” is shown here, but in the actual view, the number travels from U1 to U0, along an edge which is instantiated because there is traffic. If there was traffic from U0 to U1, the number would be shown in their appropriate viral state colors moving in the opposite direction.**

**The long slider at the bottom can roll back the clock, and then the black “<” and “>” buttons can change the viewing time by an hour and show the animation.**

**At Day 2 HR10, 2Greens are moving from U1 to U0 and 1Y+1G are just arriving at U1 from U0.**

**Clicking on the red “STOP” button stops the animation. This is useful if the traffic pattern is complex, and playing the clock back and forth allows the viewer to focus on the details of the traffic movements.**

**There will be highlighting of the node which has the highest risk (should this be largest number of infectives arriving, or largest number of infectives present, or the highest ratio of susceptibles to infectives). It would probably be good to show the names of the Universes, and their internal population and composition through right click or mouse-over, in addition to the movement of persons.**

**RELATIONSHIP TO EQUATION BASED COVID MODELS**

Classic epidemiology models of contagions use historical data (from past days) to estimate the parameters of a set of equations describing the dynamics of the epidemic, and then use the parameters to predict the future behaviour of the infection in the population.

The agent based model simulate the forward behaviour of a population using the characteristics of individual persons and the dynamics and parameters of the contagion.

In a sense, validation of each is difficult. In the first, the assumption that the past predicts the future is only really true in hindsight, and so every day brings a new challenge to validation or parameter re-evaluation. In the agent based model, the simulation is generally intended to be illustrative and easily amenable to parameters with physical analogues to demonstrate effects.

However, they should in many senses be equivalent expressions of the same large scale phenomena, of an infection affecting a population through transmission dynamics.

The key to their equivalence may be through the coefficient R0, which is defined as the number of susceptibles infected by an infective person during the time they are infectious.

In the CovidSIMVL model, R0 is easily calculated by tracking the number of susceptibles infected by all orange agents (inert), and dividing the sum by the number of orange agents to get the average R0.

Furthermore, to yield a specific value of R0, the CovidSIMVL can be used as a platform for varying different parameters to yield the same R0, and this should provide a means of assessing the effect of different mitigation strategies in different populations.

The use of multiple universes to represent the interactions of different populations affecting one another through common spaces or visitors can be quantified in CovidSIMVL, and this can therefore be a tool for defining the variations in transmission risk with different densities or duration of stay in a common space, in a rigorous and methodical manner.

**CONCLUDING REMAKRS**

The button on the top right marked “Calc R0” will produce the R0 for the entire population, given its traffic flow. To know which Universes are at most risk from what other Universes, if such correlations can be made, would be of great interest, if it can be demonstrated. For instance, the hypothesis that staff going to bars have more influence for bringing Covid into Long Term Care than visitors who go to bars can be tested.

Schools, and the effect that children have when they mingle, and go back to multigenerational homes, can be modeled in CovidSIMVL, and mitigating effects for the risks involved can be introduced and their effects seen, so that questions about school opening affects long-term care, and the extent of this link, can be assessed.

Currently, work is starting on creating the environment in R to implement goal-directed parameter optimization using CovidSIMVL as the simulation tool.

Many unknown factors and relationships are present in implementing a simulation of the SARS-2 epidemic. For example, what is the quantitative mitigating factor for masks in terms of transmission efficiency? Does it vary linearly with viral load of the infective, or an inverse square relationship?

How do observed incidence of patient characteristics convert into odds ratios, and how are they combined?

What is the extent of transmission for which we use degree of overlap as a proxy for duration of contact, and how do we combine this with the viral gradient between the two?

How does viral load affect infectivity and therefore size in the model?

There are many factors that we are starting to learn about SARS-2 but they all relate in the final analysis to how infective a person is, and how sick they get. Whether there is one or twenty factors, they combine in the human to some single variable result, and this is what CovidSIMVL attempts to simulate – a person, a set of paramaters combining to give an infective result, and operating within the dynamics of population movement and interaction.

**DESIGN OF CSV SPECIFICATION FILES**

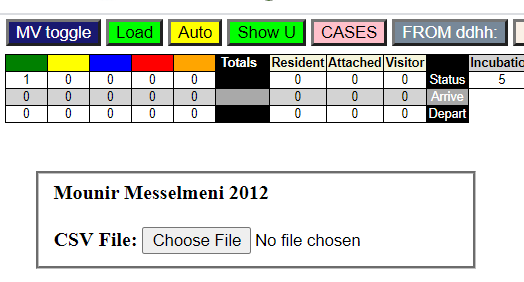
The first CSV file, which is named in the prototype as MVDATAge.csv, is a file that describes the movement of agents between universes, and the roles they have in them.

The second CSV file, which is named VL1.csv, identifies a subset of the agents in the MVDATAge.csv file as Covid cases, with viral loads, time since infection (since we are using the Temporal Dynamics and Viral Load model of Xi, He), and so on.

The program starts by requiring the use of the system to enter the population of agents corresponding to the schedule. If they do not match, there may be errors. The MVDATAge.csv persons are the for whom moves are proposed each cycle, and therefore viral transfer, viral growth, resizing are functions that are called for that population.

If the total population entered in the initial prompt is smaller than that of MVDATAge.csv, we will probably have execution errors from “undefined” array errors. If the total population is larger, the data tallies may be wrong, as the excess do not go to any universe but exist in pre-initialized state but are counted.

The program could check the match between them, but this requires error handling by the user. The program could change the CSV files to have a single multi-level parser, but that is a step for the future. As it stands, the system is bound to the FOUR steps. This is what appears after step 1 has been done.

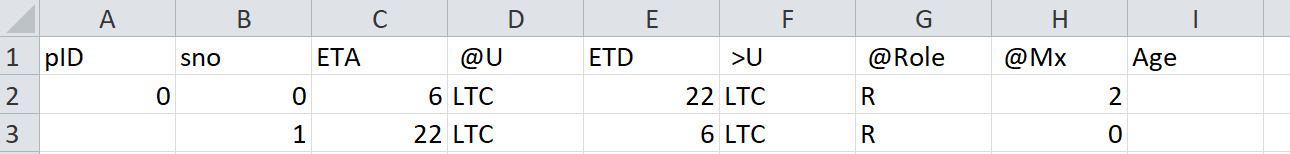


1. Input the correct population number at the startup prompt
2. Input the MVDATAge.csv file (or any other named file with data in the correct correspondence.
3. No positive correct loading is given. When control comes back to the user, click the “CASES” button and enter the VL1.csv file containing the initial cases.
4. PRESS THE LOAD BUTTON (in green). Sorry, but this is a crucial initialization step.

If the VL1.csv file is not entered, there are no infective agents, and the system will not simulate any epidemic.

MVDATAge.csv

Of course, the cvs file is created by an Excel table, saved in .csv format. The table entry looks like this:



The pID field is the unique identifier for an agent (person, patient, class of persons, etc). The simplest approach is to number them from 0 to N.

This file describes the arrival and departure times for agents with respect to different universes (the maximum number is 9, for visualization purposes). The setting, M.UCt is a program variable, not an input parameter.

Think of the model of a master schedule for a train system. It knows about TIME (Days, Hours), Passengers, Arrivals, Departures. For each Passenger, the train reservation and scheduling system has a Journey profile with a number of stops, each with an ETA (expected to arrive) and ETD.

Thus, “sno” is the Stop Number for this passenger (0). Here we show two Stops (which may be thought of as ticket stubs). The ETA for pID=0 is at 0600hrs or 6am, and “@U” indicates the place that the person is at, or will be at, at 0600.

For initialization, this first set of tickets have no corresponding departure tickets which show their arrival on line 2 above. They land from zen-space at the ETA time, into the @U universe. Here we see that it is called “LTC” (for Long-Term-Care). The “ETD” field indicates the time that this person will be departing from @U, and the field “>U” shows the Universe to which that person will be going (which in this case, is also “LTC”, which means that person=0 will be staying there.

The “role” of this person for this duration is “R” which is “Resident” which means they will stay there permanently (long term care, hospitals, submarines, cruise ships, prisons; at a different level, these could be persons who do not travel by car, train, foot, air, or ship from a city, municipality, region, county, state).

The role “A” means “Attached” and these persons would be connected on a partial basis to the universe, such as staff who come for 8-hour shifts and leave. While there, their activity level (see below) is higher than for “residents”.

The role “V” stands for “Visitors”. Persons who visit patients in hospitals, who are outpatients in clinics, who are in transit in airports, students in schools and universities, are visitors. In general, they have less mobility and mingling or exposure to others than “Attached”.

“@Mx” is the mingle factor from 1 to 10, the higher the more active. This notion of activity is expressed the following way: at each cycle, for each person, a move from current location to another location is proposed, first stochastically according to random selection from a table of values which represents a Pareto distribution (more smaller numbers than larger – I suppose in theory the number should be randomly distributed in the table rather than clustered, as they are at present). Then the mingle factor is applied, in such a way that with the maximum of 10, the distance moved is still going to be within the bounds of the universe (if the universe had no bounds then everyone would disperse further and further away).

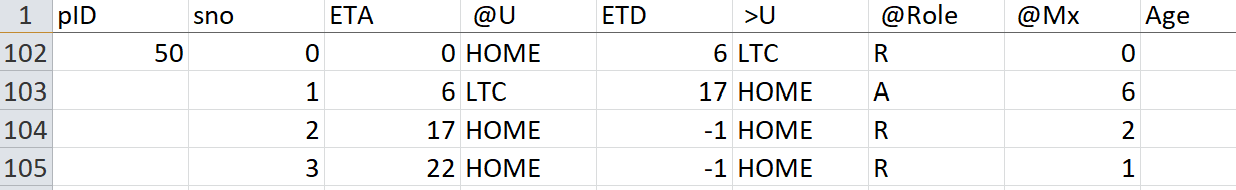
The calculation of the proposed move may also be subject to a mingle factor for the universe itself, but at present this is not a parameter but would have to be introduced through a modification of the calculation, though the field M.minglf is part of the Universe data definition already.

Essentially, the proposed move generates a pseudo-random walk (biased) for an agent, and the mingle factor increases the area covered by the random walk. Clearly a mingle factor of ZERO should cause the agent to remain in place, in which case the random walk covers an area of zero, and infection by contact cannot occur.

CovidSIMVL has the capability of defining a number of epicenters within a Universe, to which agents are stochastically drawn (but not magnetically)….thus, this feature could be used to represent persons in restaurants, coffee shops, bars, and receptions, in which most guests stay around a space, but the attendant staff mingle much more – same for students and Teaching Assistants.

The last field “Age” is for an age-group of the person. A function ageRisk(G,g) exists within the code for looking up an age risk table based on 10-year age groups, the population age distribution of BC for 2019, and the incidence distribution of age-group among Covid cases in Spain and Italy, with the risks ranging from 0.33 for age group 0 to 2.33 for age group 90+. Using the numbers is not clearly defined, so at present this calculation has been left out in the Temporal Dynamics – Viral Load model used in CovidSIMVL.

Here is another example:



Person=50 has a role of “R” in @U=HOME from ETA=00 (midnight) to 0600 with a mingle factor of zero (sleeping), and leaves at that time for LTC. At LTC (line 103), arrival is at 0600 (instantaneous transport) and there, assumes a role of “A” (could stand for Attendant), with a mingle factor of 6.

At 1700 hours, P(50) arrives at HOME and there is not departure (-1) in this ticket, but now assumes a role of “R” – resident with a mingle factor of 2. The next ticket (sno=3) at 2200 again without departing, changes the mingle factor to 1 (getting ready for bed), and then the ticket recycles to midnight (line 102) with departure at 0600 for LTC.

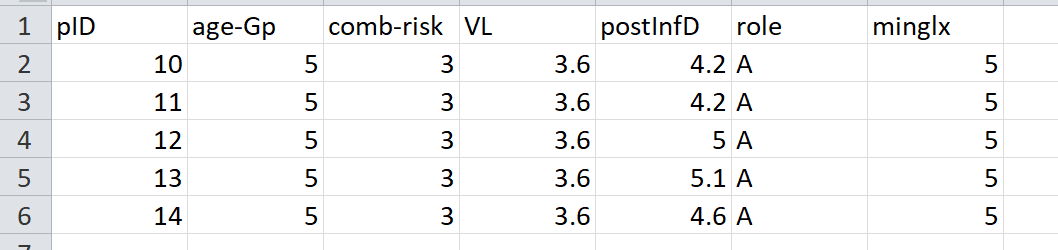
HANDLING TIME

ETA and ETD can be in the form of “DDHH” – what has been shown is just the HH. The calculation is to take this field, and use floor(ETA/100) as DD with remainder (ETA/100) as HH.

The routine called “conductor()” matches the next clock time with the times on the ticket stubs (sno) for the current DD, so that a specific schedule with Thursday being different can be constructed. Note that the default is for the schedule for the next day to use the current day if there is no explicit entry for the next day. This is how the one-day schedule becomes the default recurrent schedule if no explicit tickets for a day is found. I am not sure at this point how robust the testing is for this schedule by exception approach.

**VL1.csv File**

This file is a “Case” file and supersedes corresponding entries in the MVDATAge.csv data.



This example modifies the data for 5 persons, from P=10 to P=14. The first field, “pID”, is the unique identifier. The second is a duplicate field for age-group (and can be removed in time). The third is a combined-risk factor, to cover innate risk factors such as obesity, diabetes, asthma, vaping, hypertension, serious heart disease, chronic kidney disease, perhaps race. Rather than entering these in the simulation program, and requiring it to combine these risk factors, I have made the decision that this should be done better externally by the model creator, and combined into a single factor which affects the size (or viral growth? Or mobility?) of the person.

The “VL” field stands for Viral Load according to the Xi,He paper, using the CT times as score. The postInfD is important – this is the number of days after infection. For example, a newly infected person in the simulation could be given a number here of 1 but this would require the simulation to run for 1.9 days before the person becomes infective.

A creative example of this specification is to give the agent a mingle of 0, and a postInfD of 12.2d with a high viral load, say of 6. This then places the agent with 1 day left for infectivity, and immobility – exactly what a door knob would be.

These CSV files should be considered evolutionary, and their evolvement is inevitable.