Modeling and Visualization

What is modeling?

Modeling is actually something that people do all of the time without realizing it. What is modeling? Modeling is a method to simplify a complex system in order to better understand that system or to make decisions about system behavior and management of the system. Models are an abstraction of reality designed to allow their creators to gain insights into how systems work, how to manage them, and how to make informed decisions. In science, they are a way of conducting inquiries about complex systems to address these questions.

To illustrate, let's start with an everyday example. Let's say you need to decide the best route to take to work. To make this decision you will need to formulate at least one objective for your trip. Here are some of the possible objectives.

- Minimize the amount of time it takes to get there
- Avoid traffic congestion (You hate traffic!)
- Find a route that excludes freeways (You have an old clunker of a car)
- Plot a path between your house and work to make sure you travel by the same spot every day (You could think of this as any stop along the way - for coffee, breakfast, to follow construction development on a friend's new house, etc.)

If we focus just on the first objective on the above list, we need to decide what will affect that objective. To do this, we must create an abstraction of what is a complex system and set of decisions. We have a finite but potentially large number of routes we can take. Each section of

roadway has a particular speed limit. Each intersection might cause us to slow down because of a stop sign or a traffic light. The amount of traffic on each section of roadway will affect how fast we can drive. This varies dynamically by time of day -- at peak rush hour there is more traffic and it moves slower. What time of day is your work trip compared to this peak? Other things that can slow us down are construction, the occasional accident, or changes in the weather. If the weather is bad, we might change the path we take. In snow, the main streets are often plowed and salted and thus move much faster than the side streets we might take on other days.

Based on our own experiences, we abstract this complex reality to create a mental model of what we perceive to be the best route that meets our objective. Every day we take this route and this gives us real data that may or may not match our original expectations. For example, we may have omitted information about intersections in our original model, but find that one particular traffic light (where we have to make a left turn) becomes a major problem during congested
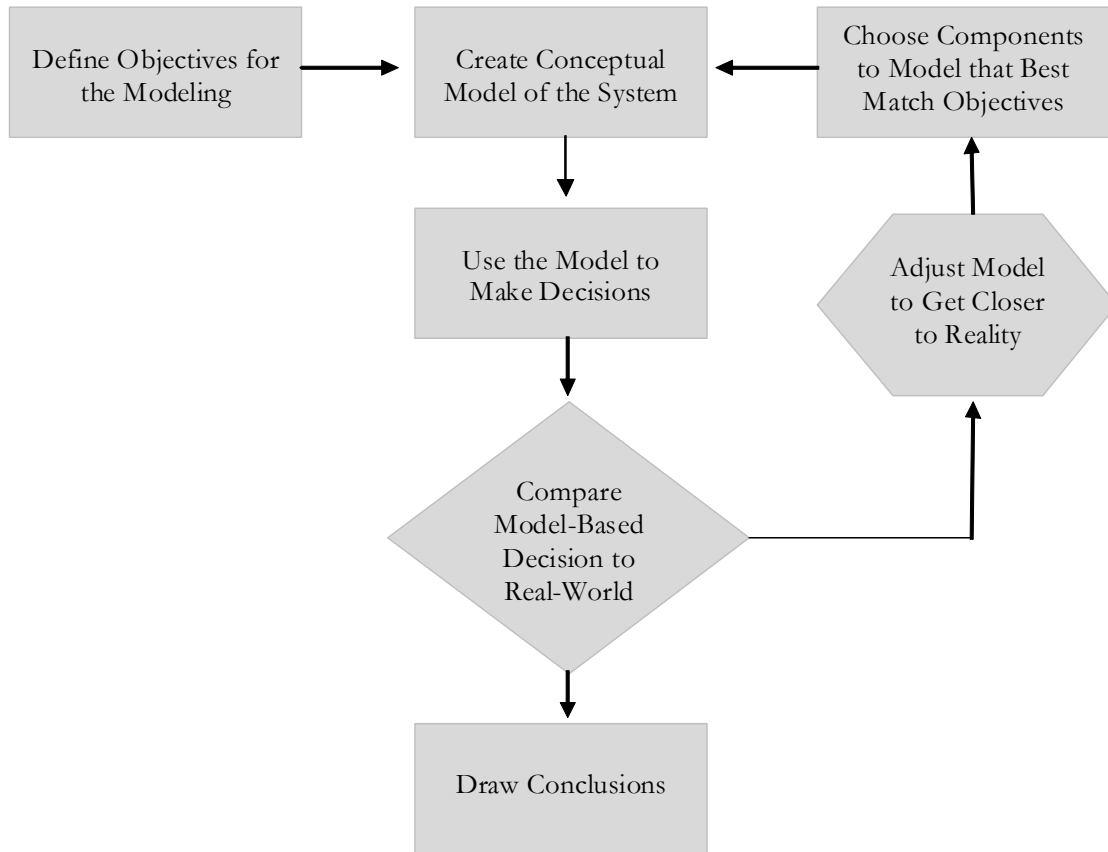
times; it may take three or four light changes to get around the corner. In essence, we are comparing real data to our model, and adjusting the model and our route accordingly until they get us closer to our objectives.

Using your experience with traffic, try to conceptualize a model that forecasts the time it will take you to go from your home to work or school. Given what you know about the possible routes, what are the variables that should be considered in your model? How will you represent each variable? To simplify your model, you may wish to make some of the conditions constants or **parameters**. A parameter is a variable that is given a specific value to represent one set of conditions for the model. In reality those conditions might vary but not enough to significantly change the model results. For example, every stop sign you stop at may take a different amount of time depending upon how many other cars are approaching the same intersection. However, you may choose to create a parameter that uses an average amount of stopping time to represent the range of conditions. Make a list of all of the possible variables, those you will choose to make into parameters, and how you might estimate the value of each (e.g. go out and measure the time it takes for an event, use Mapquest to calculate the distance along each roadway link and the travel time, etc.).

If we formalize this process it might look like the figure below.

The Modeling and Inquiry Process

Figure 2: The Modeling and Inquiry Process

```
┌──────────────────┐      ┌──────────────────┐      ┌──────────────────────┐
│ Define Objectives │─────▶│ Create Conceptual│◀─────│ Choose Components     │
│ for the Modeling  │      │ Model of the     │      │ to Model that Best    │
│                   │      │ System           │      │ Match Objectives      │
└──────────────────┘      └──────────────────┘      └──────────────────────┘
                                   │                           ▲
                                   ▼                           │
                          ┌──────────────────┐        ╱───────────────╲
                          │ Use the Model to │       │  Adjust Model   │
                          │ Make Decisions   │       │  to Get Closer  │
                          │                  │       │  to Reality     │
                          └──────────────────┘        ╲───────────────╱
                                   │                           ▲
                                   ▼                           │
                             ╱───────────╲                     │
                            │  Compare    │                    │
                            │ Model-Based │───────────────────▶│
                            │ Decision to │
                            │ Real-World  │
                             ╲───────────╱
                                   │
                                   ▼
                          ┌──────────────────┐
                          │ Draw Conclusions │
                          └──────────────────┘
```

As we implement the model, we need to go through several steps to ensure that the model works correctly. First, we need to **verify** that the mathematical representation of the system is correct and that the computer program we create to make the calculations is giving us the expected answers. Second we can adjust the model to more closely reflect reality or **validate** its accuracy. To do this, we will need experimental data that allows us to compare the model results with the experimental results. We must also decide the acceptable level of accuracy given the purpose of the model. In our traffic model, if we are off in our estimate of travel time by a minute or two, that is probably inconsequential. If it were 15 minutes and that would potentially make us late for work, it might be more important to improve model accuracy.

In order to implement this first example, we have created a MATLAB representation of a simple model of a trip to work. You will need to have MATLAB installed on your computer to use this model. When you are ready open up the traffic1.m MATLAB model.

At the top of the code, you will see the definition of the major variables that are used to estimate total travel time. This is followed by a list of parameters that are used to estimate the amount of time it will take to traverse the two alternate routes to work. For each leg of the trip on a local street, the travel time is the distance divided by the average travel speed. Time is increased if

you get to a stop sign and here is represented as a constant parameter of 30 seconds. The time to traverse a section of highway or a major street is the same as for a local street except that the average speed will be higher. If you get to a traffic light, the time will actually vary depending on the probability that you will hit a red light and the average amount of cycle time for the light to turn green. Finally, our model adds time if we need to make a left turn at a major intersection as this will delay us further. Assuming there is a green turn arrow, we need to know its cycle time and the number of cars that can make the turn per cycle time.

The next part of the program shows three route options based on given/assumed data. The model considers two routes - a freeway route and a local road route. The freeway route is considered for two conditions - with and without congestion. The path associated with each option is
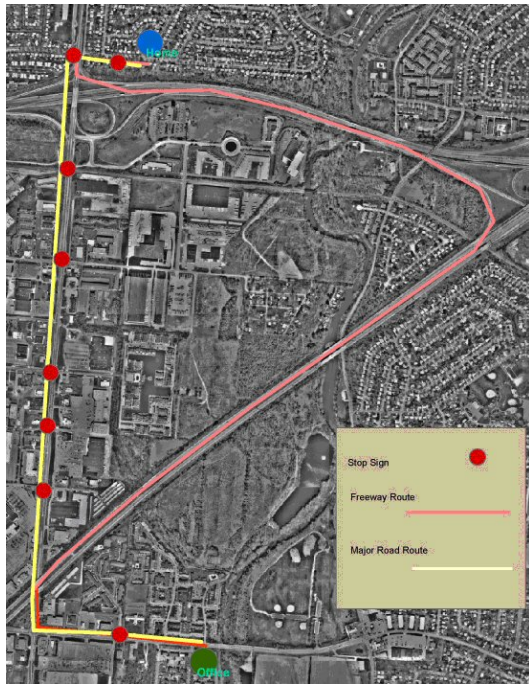
described by calculating values for each part of the trip and placing the results into a vector. For the freeway route one takes the local streets to get to the freeway and then travels 2.5 miles to the freeway exit. From there, one has to negotiate a left turn and travel several blocks to the destination. For the local street option you use a major road and travel 1.75 miles and face the possibility to be stopped at 5 traffic lights. The difference between congested and normal freeway is the difference in speed (55 mph uncongested freeway, 20 mph congested freeway). The alternate routes are illustrated in the picture.

Take a few minutes to look over the formulas in the program and make sure you understand how each segment of the trip is used to calculate the total time for that segment. You in fact must verify that the mathematics of the equations correctly calculate the travel time given the assumptions that are made about the relationships. If you are confused about any of the calculations, you can post a note to the discussion area. Compare what is simulated in the model to your conceptual model of travel time. Which of the variables you thought might be important are left out of this model? Under what circumstances would the model probably fail to accurately forecast travel time?

Once you are sure you understand the model components, take some time to use the model to simulate alternatives. Simulation uses alternative sets of conditions to make a series of forecasts, giving the analyst insights into model behavior and to make a rational decision about what to do. For this example, use the program to simulate the differences in travel time if the probability of hitting a red light changes. We are going to change the appropriate parameter from .01 to .5 in .01 increments. The program will recalculate the travel times. Do this for both ends of the sequence and record the results. Testing these types of variations are called **sensitivity testing** for the model. The question is how sensitive is the model output (the time to get to work) to variations in the constants that are used as inputs. One way to measure sensitivity is to quantify the percent change in output per unit change in an input parameter. So in the case above, you

should answer the question: for each 1% change in the probability of stopping at a traffic light how much is the average time to work increased or decreased? Is the model sensitive to this variable?

To assist with the analysis, we have provided another version of the program that automatically calculates the values throughout the range and creates two graphs of the results. This is done by making the main content of the program a function embedded in traffic1b.m. The function calculates the values for the three scenarios every time it is called. The function is called in the program run_traffic.m. Open that program and look at the code. There is a loop that calculates the traffic for the range of red light probabilities and makes two graphs – one which is drawn after all values are calculated and one which plots points as they are calculated.

Run the run_traffic.m program and use the results to answer the question about which route is best to take under which conditions. Describe the general nature of the model and your findings in a report which includes relevant data and graphs.

Test the sensitivity of a second parameter – the average traffic speed on major roads assuming that they also become congested and slow down during rush hour. In order to compare these runs with the original runs, add a title to each of the graphs that indicates which major road traffic average speed is being simulated. How does a slow down on the major streets impact the relationships between scenarios in the model?

Extended Exercise

Traffic engineers that have studied highway congestion have used empirical data to create a more refined model of the traffic flow associated with congestion. The Business Process Reengineering (BPR) method uses this formula (Small and Evans, 1989):

$$TT = \alpha + \gamma \left(\frac{V}{C}\right)^k$$

where:

- TT is the travel time per mile, in minutes on a given road
- V is traffic, measured in passenger-car-equivalents (pces) past a given point, per hour.
A quick-and-dirty way to compute this, given that you know the composition of the traffic stream on a relatively level terrain is

$$V = (\#cars) + (1.6 \times \#buses) + (2 \times trucks)$$

For our example, we will assume that all of the traffic on our section of the highway are usually cars.

- C is the road capacity = lanes X (capacity per lane in pce's per lane per hour). Typical values here are lane capacity = 1800 - 2000 for an urban expressway and 1000-1100 for an urban arterial. We are simulating an expressway with two lanes.
- k is a parameter, usually taken to be 4.0 for an urban expressway and 2.5 for an urban arterial
- α: this parameter represents your estimate of the travel time per mile on the road in question when there is no congestion (the free-flow travel time), in minutes per mile. For example, if on an expressway you think that traffic will move at 65 mph when there is no congestion, then

α= 60/65. = 0.92308

- $\alpha + \gamma$ is your estimate of the travel time when the V/C ratio is 1, often thought of as the travel time when the road is operating at its design capacity. For example, if you think that traffic will move at 25 mph at capacity, then $\alpha + \gamma$ = 60/25. = 2.4 and γ =2.4 – 0.92308 = 1.4769. We will assume this to be the case for our example.

Create a new part of the program which calculates TT for car traffic volumes starting at 400 cars and going up by 400 cars until you reach 4000 cars. Be sure to calculate the travel time for our example as 2.5 times (2.5 miles) the TT values. Plot the input volumes against TT on an X-Y graph. What form does this relationship take?

Write a brief description of what you found using the model being sure to answer these questions:
1. What did the initial comparisons of timing for the scenarios shown originally tell you about the travel time using these two routes with and without a congested freeway?
2. What happened to those timings when you changed the probability of hitting a red light? What is the sensitivity of the model to the values of these parameters?
3. What is the best route to take if there is an accident on the freeway? What is the form of this mathematical relationship when you plotted the values and why would you expect this to be true given the model's mathematical composition?
4. When you used the BPR model to represent traffic flow on the freeway, how do the travel times compare to the estimate in the simpler model? How does the travel time change as congestion increases? Why would you expect this form based on the mathematical representation in the BPR model? How would this formulation of the model change your decision about which route to take as congestion increased?
5. What other factors from your conceptual model impact traffic that are not considered in the current model? How might you represent each in the model and what impacts will they probably have on travel time?

Create a short report answering these questions, inserting tables, graphs, or numbers where appropriate, and upload your report file to the assignment area for the traffic model (section 4.4 of this module).

References
Kenneth A. Small, Clifford Winston and Carol A. Evans, "Road Work", Brookings, 1989. The relevant discussion is p. 103 and table 6-1.

Kenneth A. Small, "Urban Transportation Economics", Harwood Academic Publishers, 1992, p. 70.

Simple Model of a Trip to Work

Model Parameters Definitions	Description			
Time to traverse block on local street	Distance / avg. speed			
Stop sign average stop time	30 second delay			
Time to traverse highway section	Distance / avg. speed			
Traffic light stopping time	Probability of green x avg. red			
Left turn on major road	(total cars turning/cars per green arrow) x light cycle			
Alternate Routes		**Model Time**		
Freeway Route	Local Streets 1	Freeway Route No Congestion	Freeway Route Congestion	Local 1
Home	Home			
2 Blocks	2 Blocks	0.909090909	0.909090909	0.90909091
Stop Sign	Stop Sign	0.5	0.5	0.5
2 Block to Freeway	2 Blocks to Freeway	0.909090909	0.909090909	0.90909091
2.5 Mile on Freeway	1.75 Miles on Major Street	2.727272727	7.5	3
Left Turn on Major Road	5 Traffic Lights	4	4	5
1 Block	Left Turn on Major Road	0.454545455	0.454545455	4
Traffic Light	1 Block	1	1	0.45454545
2 Blocks to Work	Traffic Light	0.909090909	0.909090909	1
	2 Blocks to work			0.90909091
Total Time (minutes)		11.40909091	16.18181818	16.6818182

Parameters Values Used in the Model Parameter	Value	Units
Average City Block Length		1000 Feet
Local Street Speed		25 mph
Major Street Speed		35 mph
Normal Freeway Speed		55 mph
Congested Freeway Speed		20 mph
Probability of Red Light at Major Intersection		50%
Average Red Light Time		2 Minutes
Cars per Green Arrow on Left Turn		6
Left Turn Light Cycle		4 Minutes
Conversion Miles to Feet		5280
Conversion Hours to Minutes		60
Cars Turning Left During Peak		12
Cars Turning Left Off Peak		4

Grading Rubric

Assignment Requirement	Suggested % of Total Grade
Discussion of benefits of modeling, importance of modeling, and example of a modeling and simulation project in current science and engineering work	10%
Correct results and discussion on initial linear model of traffic	15%
Sensitivity testing of the red light probability parameter	15%
Impact of an accident on travel times; linearity of relationships	15%
BRP model correctly implemented in spreadsheet	20%
Discussion of travel times with BPR model; comparison to linear form; discussion of non-linear function	15%
Discussion of other traffic factors left out of this simple model	10%

Rubric to Assess Student Work in Computational Science

	Unacceptable 0	Below Average 1	Average 2	Superior 3	Score
Writing Mechanics	Many spelling and grammar errors	Several spelling and grammar errors	Few spelling and grammar errors	No spelling and grammar errors	
Writing Quality	Incompetent, not college level	Marginal, college level, but below course level	Competent, appropriate for course level	Sophisticated, graduate-level	
Completeness	Missing much information	Missing some information	All basic information is included	Goes beyond assignment's minimum requirements	
Mathematics Correctness	Many errors	Few errors	No errors	No errors and shows depth of understanding	
Science Correctness	Many errors	Few errors	No errors	No errors and shows depth of understanding	
Computing Correctness	Many errors	Few errors	No errors	No errors and shows depth of understanding	
Logic and Flow	No logic to the flow of information	Awkward flow of ideas	Logical flow of ideas	Attempts to tell a story	
Use of Sources (if appropriate)	Has less than 75% of sources	Has at least 75% of sources or not all are cited	Has minimum number of sources	Has more than minimum number of sources	
Interdisciplinary (Math, Science, Computing)	Missing two components	Missing one component	Attempts to incorporate all components	Effectively communicates all components	
Computing Technology	Technology not used	Technology level below assignment requirements	Technology level appropriate for assignment	Technology level higher than assignment requirements	
Creativity	Does not demonstrate even "textbook" or lecture thinking about project	Does not go beyond "textbook" or lecture thinking about project	Demonstrates novel thinking about project	Effectively and convincingly demonstrates novel thinking about project	

Rubric based on Karkowski, Capital University, 2006.