

# Decision Aid Methodologies In Transportation

## Lecture 1: Introduction

---

### **Operations Research and its applications on decision making in transportation systems**

Shadi SHARIF AZADEH

Transport and Mobility Laboratory TRANSP-OR  
École Polytechnique Fédérale de Lausanne EPFL

# Overview

---

- Some examples on the success and failure of decision making policies
- Decision science of the 21<sup>st</sup> century
- Data uncensoring methods
- Solving an example

## Some examples of wrong decision makings

---



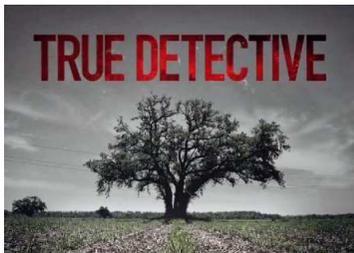
LG Says Google Underestimated Nexus 4 Demand. Google severely under-estimated the demand for their Nexus 4 smartphone, leading to the shortage facing most eager customers across Europe and the United States. They had to boost the production in order to respond to the customers of T-mobile who had a contract with LG and Google.

Ref: Karl Bode Jan 22, 2013

# Some examples of wrong decision makings

---

**HBO GO**<sup>SM</sup>



HBO GO is the successor to HBO on Broadband, originally launched in January 2008, consisting of 400 hours of movies, specials and original series (including 130 movie titles that rotate monthly) that could be downloaded to computers, at no extra charge for HBO subscribers. Meltdowns in HBO Go happen usually on Sundays and it affects the stock market. That's only because HBO is now the prime destination for some of the greatest TV shows existing today. But when you see their server crash more than once, you have to wonder whether the demand for certain TV shows is being continually underestimated.

Ref: Greg Brian Apr 10, 2014

## Some examples of wrong decision makings

---



Disney has admitted to underestimating the popularity of the film, which has so far reaped \$1.2 billion at the international box office. Shortage in Frozen merchandise has triggered an inflated online black market with desperate parents willing to shell out big money for the popular Disney toys. Parents are now spending hundreds of dollars to import merchandise toys from the Disney film, with dolls being sold on eBay for as much as \$1,000 and dress up costumes ranging from \$174 to \$530.

Ref: Emily Crane, 24 May 2014

# Some examples of wrong decision makings



- Office of Rail Regulation found 115,000 people were affected by problems.
- Paddington and King's Cross were to reopen on December 27 after works.
- But Paddington was closed all morning and King's Cross all day.
- Paddington safety work which should have taken two hours took ten.
- People faced 'widespread confusion, frustration, discomfort and anxiety'

*“Bad planning, the failure of contingency arrangements and breakdown in communication resulted in thousands of people waiting for hours”*

Louise Ellman, chairman of the transport select committee

<http://www.dailymail.co.uk/news/article-2950566/Passengers-really-let-train-chaos-ruined-Christmas-says-damning-report-Network->

# Some examples of difficult decision making



The territory covers 6.843 square kilometres and shares a land border with Spain to the north. The Gibraltar Airport is 487 meters from the city, the shortest commute of any major airport in the world (1,680 m length of runway). One would naturally ask the question how difficult it is to operate and land aircrafts when the airport is so close to the city. British Gibraltar has very little area, and the important airport runway takes up a major portion of land.

## Some examples of wrong decision makings

---



Taxpayer-supported University Medical Center (UMC) in Nevada has been forced to borrow \$45 million in just four months to cover a flood of new Medicaid patients signing up via Nevada's expansion of the program through the Affordable Care Act "Obamacare". The reason, according to the *Sun*, is that the state underestimated the number of new enrollees through the expansion of Medicaid from Obamacare.

Ref: Michael Chamberlain Apr 28, 2014

# Data Driven Decision Making

---

Minority Report (2002)



1946-Now



2054

# Data Driven Decision Making



1955-2011



2014

# Data

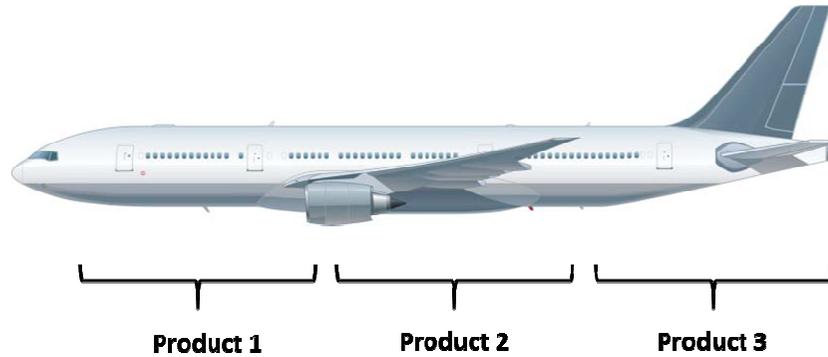
---

No good comes without a price.

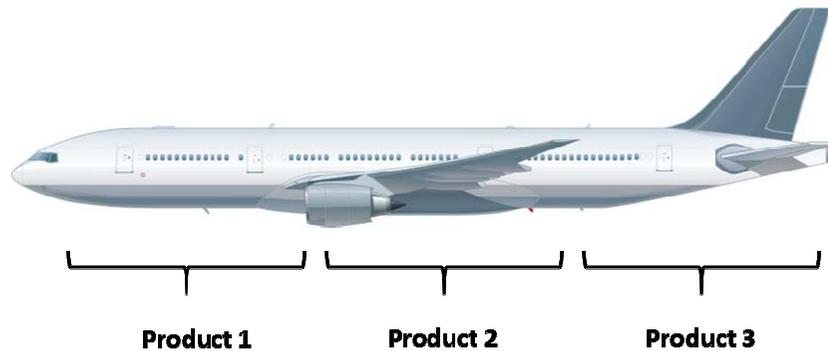
- Curse of dimensionality
- Missing information and Data censorship

# Uncensoring data

Monday 8:00 a.m.

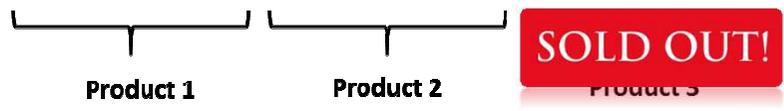


Monday 10:00 a.m.



# Uncensoring data

Monday 8:00 a.m.

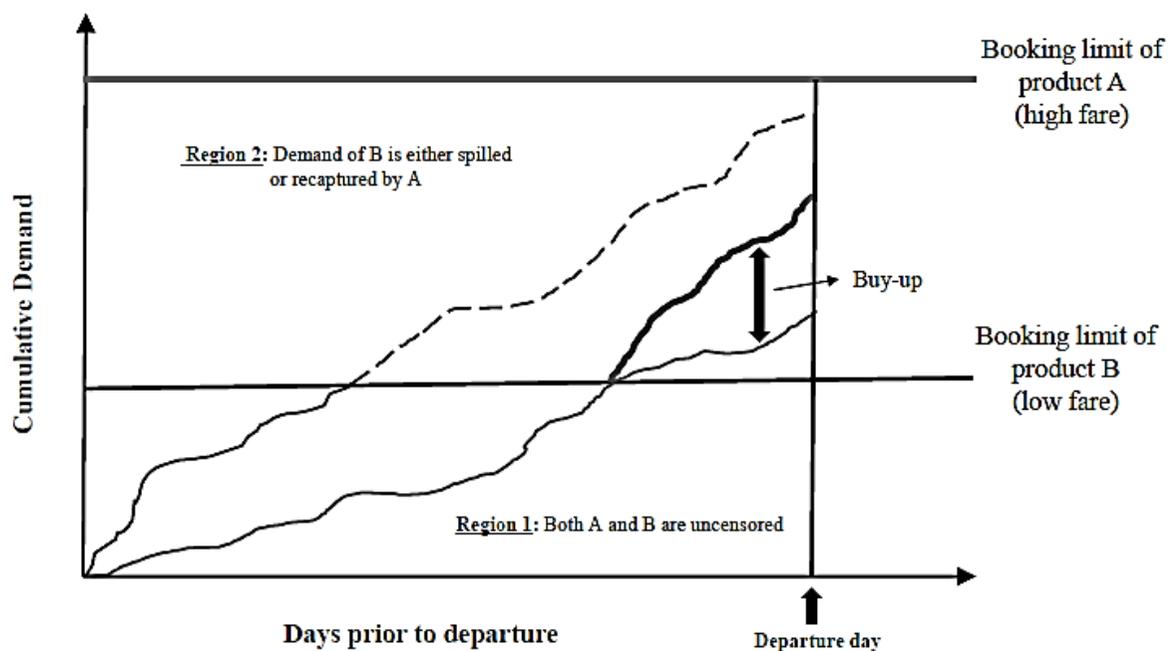


Monday 10:00 a.m.



Spill

# Importance of Uncensoring data



Underestimating demand by 12.5% to 25% can result in a loss of revenue from 1% to 3%, which is Significant.

Uncensored demand of A where A and B are independent ———

Uncensored demand of A when demand of B is recaptured by A (buy-up) ———

Uncensored demand of B ———

Censored demand of B - - - - -

Weatherford and Belobaba (2002).

## Methods to Uncensor Data – Basic Methods

### Basic methods:

Product	A	B	C	D
<b>General information</b>				
Availability	1	1	0	1
Booking limit	8	6	0	23
Actual demand	15	0	9	22
<b>Uncensored demand by direct observation</b>				
Registered demand (Direct observation)	12	0	0	22
Direct observation	12	0	0	22
<b>Uncensored demand by :</b>				
Registered demand (Other three methods)	8	0	0	22
Ignoring censorship	8	0	0	22
Discarding censorship	-	0	-	22
Mean imputation	10	0	0	22

# Methods to Uncensor Data – Statistical Methods

---

## Statistical methods:

### 1. **Historical booking models**

Time series Box et al. (2011)

Exponential smoothing Hyndman et al. (2008)

Linear regression Lee (1990)

### 2. **Advanced booking models**

Pickup methods Gorin (2000); Mishra (2003); Zakhary et al. (2008)

### 3. **Combined models**

Weighted average method Wickham (1995)

Distribution based demand Popescu et al. (2012); Eren and Maglaras (2009)

Neural networks Weatherford et al. (2003); Sharif Azadeh et al. (2012)

# Methods to Uncensor Data – Statistical Methods

## Time series

Despite their relatively simple mathematical structure, they are rich enough to embody a wide range of data features. For one, the ARIMA model comprises autoregressive and moving average components (Box et al, 2011).

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} \\ + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

## Exponential smoothing

On the basis of data observed up to time  $t-1$ , Simple exponential smoothing adjusts the next value through the formula

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)$$

the parameter  $\alpha$  lies between 0 (no adjustment) and 1 ('strong' adjustment). This method, which relies on a weighted average of the most recent observations (Hyndman et al, 2008), is not recommended for the analysis of time series characterized by a large number of null values and a high variability among the non-zero data.

# Methods to Uncensor Data – Statistical Methods

---

## Regression

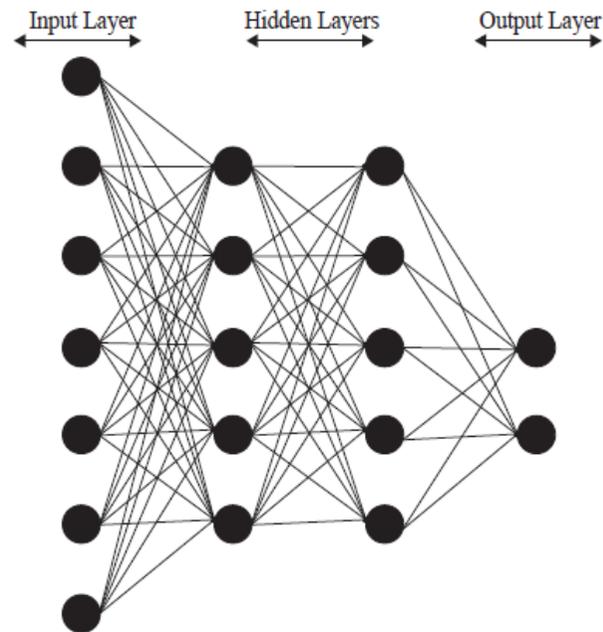
Linear regression assumes a linear trend of registered bookings in successive time periods, the key issue being to properly select the number and nature of the descriptive variables entering the model. The parameters of the regression are usually estimated via least squares. For a case involving two descriptive variables over two successive booking intervals, we have that

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2}$$

# Methods to Uncensor Data – Statistical Methods

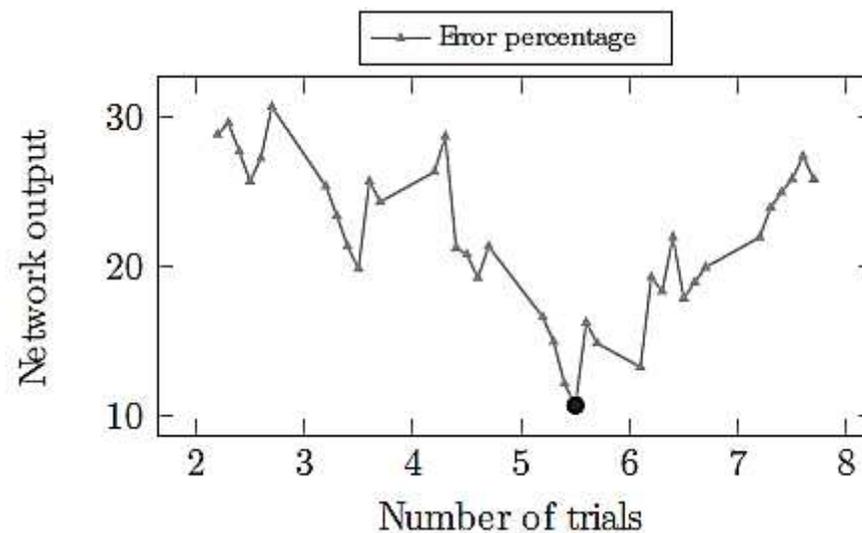
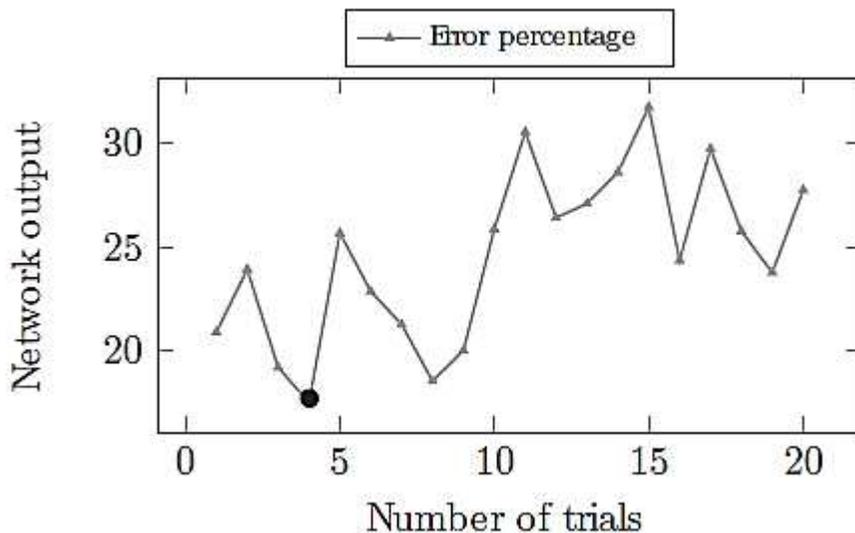
## Neural networks

Supervised learning neural networks are able to process large and complex data sets. A neural network comprises an input layer, one or several hidden layers and an output layer. In the “training phase” one iteratively adjusts each weight until the difference between expected and actual data falls below a predefined threshold value. Following this phase, the network is used to predict future values from a data set that should not differ too widely from the training set.



# Methods to Uncensor Data – Statistical Methods

- Pre-processing (outliers, choice of activation function, normalization)
- Structure of the network
- Choice of the learning algorithm (back-propagation)
- Sigmoid function that exhibits a balance between linear and nonlinear behavior
- A line search method of finding a local minima (regularization, steepest descent)
- Adaptive learning to boost the model performance



# Methods to Uncensor Data – Statistical Methods

---

## Distribution based models

In Distribution-based models, it is assumed that the statistical distribution underlying the process (usually Normal or Gamma) is known, and that its parameters (mean, variance and so on) are estimated based on historical data. Alongside the Normal or Gamma assumptions, Brummer et al (1988) has considered log-normal distributions, while Logistic, Gamma, Weibull, Exponential and Poisson distributions have been advocated (see Kaplan and Meier, 1958; ZF Li and Hoon Oum, 2000; Swan, 2002; Guo et al, 2011; Eren and Maglaras, 2009; Huh et al, 2011; Popescu et al, 2013).

# Methods to Uncensor Data

## Expectation Maximization

After its introduction in the late 1990s by Salch (1997), the two-stage EM process has quickly become one of the most popular unconstraining methods. In the first step, E-step, unobserved data is replaced by its average observed data. In the subsequent M-step, the parameters of distribution (mean and variance) are estimated via maximum likelihood. The first step is then repeated and the fixed-point process is halted when no significant progress is observed. In this setting, seasonality is usually ignored.

Initialisation: Estimate  $\mu$  and  $\sigma$ , based on  $N_2$  uncensored observed data:

$$\mu = \frac{1}{N_2} \sum_{i=N_1+1}^{N_1+N_2} Y_i$$

$$\sigma = \sqrt{\frac{1}{N_2} \sum_{i=N_1+1}^{N_1+N_2} (Y_i - \mu^{(0)})^2}$$

# Methods to Uncensor Data – Optimization

---

## E- Step:

For a given number  $C$  of constrained observations, the first and second moments of the censored data required to form the log likelihood function are estimated according to the formula: iteratively to replace the missing data to form the complete log-likelihood function where  $C$  represents registered constrained observation.

$$\hat{Y}_i^{(+)} = E[Y \mid Y > C, Y \sim N(\mu, \sigma)]$$

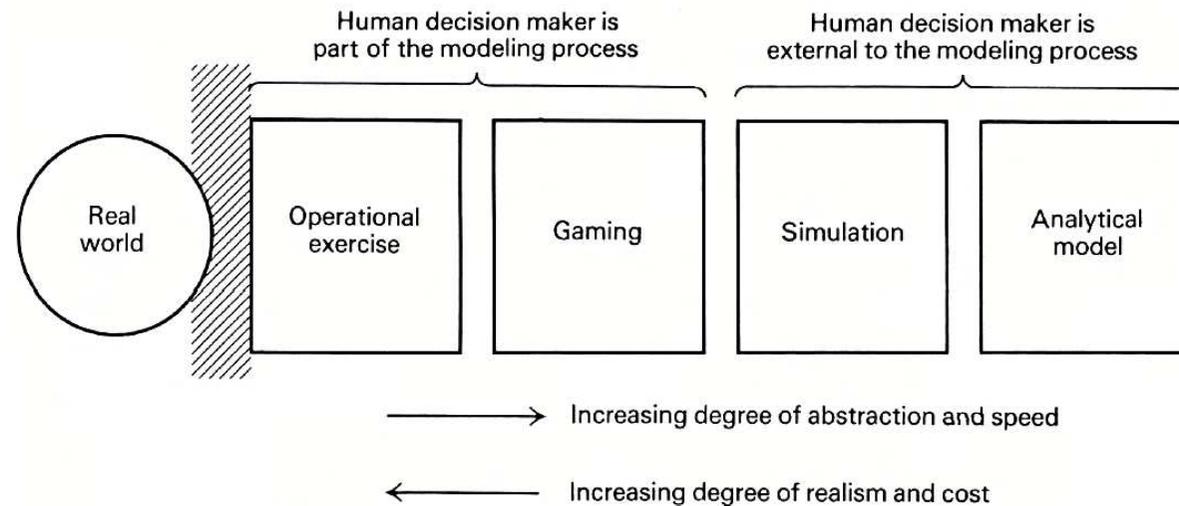
$$(\hat{Y}_i^2)^+ = E[Y^2 \mid Y > C, Y \sim N(\mu, \sigma)]$$

M-Step: Maximize the log-likelihood function with respect to  $\mu$  and  $\sigma$  to obtain  $\mu_+$  and  $\sigma_+$ .

Stopping criterion: Repeat steps E and M until the difference between successive iterates is less than some predetermined threshold value  $\delta$ .

# Problem solving methods

## Model Classification

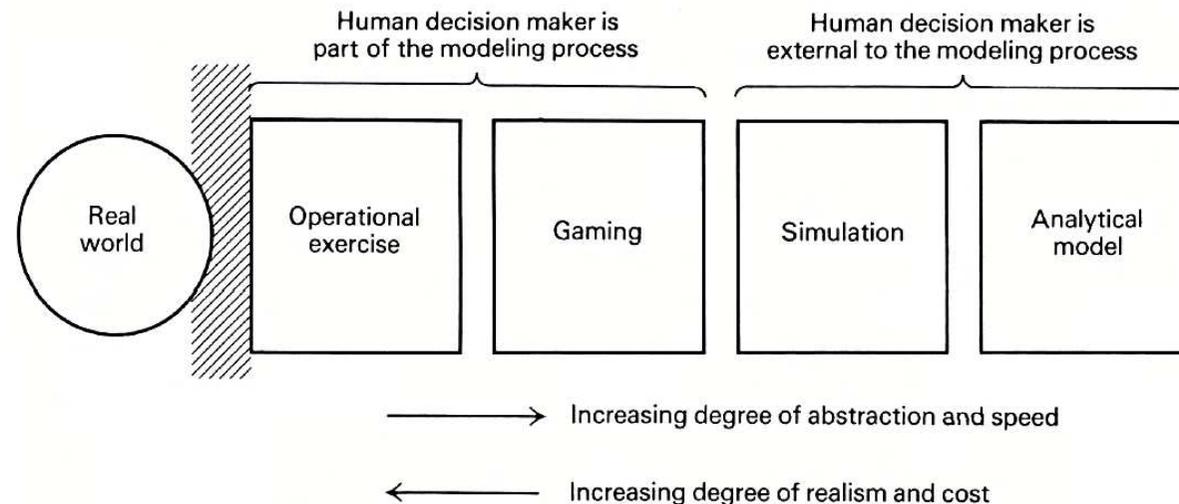


## Operational Exercise

- This modeling approach operates directly with the real environment in which the decision under study is going to take place.
- The method is expensive to implement.
- It is impossible to exhaustively analyze the alternatives available to the decision-maker → severe sub-optimization

# Problem solving methods

## Model Classification

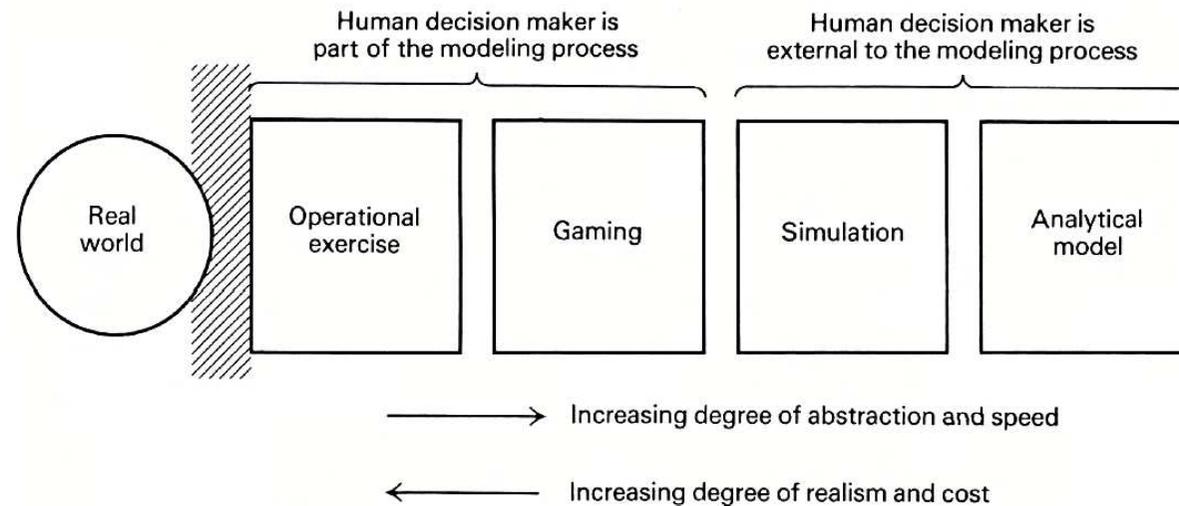


## Gaming (lab experiment)

- A model is constructed that is an abstract and simplified representation of the real environment. This model is simply a device to allow the decision-maker to test the performance of the various alternatives that seem worthwhile to pursue.
- Certainly, we lose some degree of realism in our modeling approach ; however, the cost of processing each alternative is reduced, and the speed of measuring the performance of each alternative is increased.

# Problem solving methods

## Model Classification

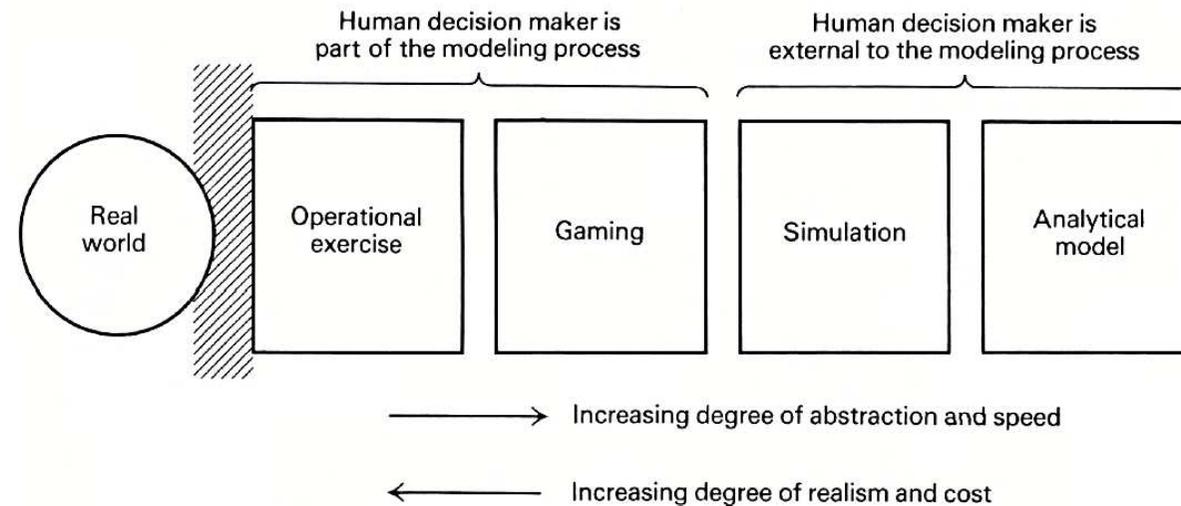


## Simulation

- Similar to gaming models except that all human decision-makers are removed
- Like operational exercises and gaming, simulation models do not generate alternatives to improve the system.
- They are useful only to assess the performance of alternatives identified previously by the decision-maker. It is a form of computer programs, where logical arithmetic operations are performed in a prearranged sequence. (e.g. traffic simulators)

# Problem solving methods

## Model Classification



## **Analytical Model**

- The problem is represented in mathematical terms, normally by means of objective. It is subject to a set of mathematical constraints that portray the conditions under which the decisions have to be made.
- The model computes an optimal solution, i.e., one that satisfies all the constraints and gives the best possible value of the objective function.
- Least expensive, highest degree of simplification in model representation.

# Mathematical Programming

---

- 1) Linear Programming (Best developed and easy to solve approach)  
U.S. Air Force known as Project SCOOP 1947 (Scientific Computation Of Optimal Programs), developed the simplex method for solving the general linear-programming problem. (George Dantzing)
- 2) Integer Programming (could be easy or very difficult to solve)  
Network (Logistics)-shortest path Dijkstra Algorithm  
Scheduling problems (cross-docking)
- 3) Non-linear Programming (Very difficult to solve)  
Traffic control (nonlinear cost on arcs)
- 4) Dynamic Programming (Very difficult to solve)  
Time depended problems (route choice)

# Mathematical Programming

## Stages of formulation, solution, and implementation

### Step 1: Formulating the model.

- Selection of a Time Horizon (Week, Month, dividing time)
- Selection of Decision Variables and Parameters
- Definition of the Constraints
- Selection of the Objective Function

### Step 2: Gathering the data.

Having defined the model, we must collect the data required to define the parameters of the problem.

The data involves:

- the objective-function coefficients,
- the constraint coefficients
- the right-hand side

### Step 3: Obtaining an optimal solution.

- Finding optimal solution is not an easy task
- Example. Combinatorial problems (TSP), Matrix Size (# of variables)

# Mathematical Programming

---

## Step 4: Applying sensitivity analysis.

- Data uncertainty and input errors

## Step 5: Testing the solution.

- The solution should be tested fully to ensure that the model clearly represents the real situation.

# Mathematical Programming

Product: 1000 lbs of casting

Manganese  $\geq 0.45\%$

$3.25\% \leq \text{Silicon} \leq 5.50\%$

Casting Price 0.45/lbs

Melting cost 0.005\$/lbs

**Question:** Out of what inputs should the foundry produce the castings in order to maximize profits?

Input	A	B	C	Manganese
Silicon	4%	1%	0.60%	0%
Manganese	0.45%	0.50%	0.40%	100%
\$/1000lbs	21\$	25\$	15\$	8000\$

Choice of input for maximum profit

# Mathematical Programming

---

## Decision variables:

$x_1$  = # of 1000 lbs of pig iron A

$x_2$  = # of 1000 lbs of pig iron B

$x_3$  = # of 1000 lbs of pig iron C

$x_4$  = # of lbs of pure manganese

## Objective function:

Max Profit (en \$)

Max Revenu – Cost

Max  $0.45 \times 1000 - (26x_1 + 30x_2 + 20x_3 + 8x_4)$

Equivalent to

Min  $26x_1 + 30x_2 + 20x_3 + 8x_4$

# Mathematical Programming

---

Total production (lbs)

We want exactly 1000 lbs of casting

$$1000x_1 + 1000x_2 + 1000x_3 + x_4 = 1000$$

Manganese restriction

At least 4.5 lbs manganese in 1000 lbs of casting

$$4.5x_1 + 5.0x_2 + 4.0x_3 + x_4 \geq 4.5$$

Silicon restriction

$$32.5 \leq 40x_1 + 10x_2 + 6x_3 \leq 55.0$$

All values should be positive

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$$

# Mathematical Programming

---

## References:

- ***Applied Mathematical Programming*** by Bradley, Hax, and Magnanti (Addison-Wesley, 1977) Chapter 1,5
- ***Introduction to Operations Research***, Hillier, Liberman 7 Companies, 2001
- ***A taxonomy of demand uncensoring methods in revenue management***, Sh. Sharif Azadeh, P. Marcotte, G. Savard
- ***Railway demand forecasting in revenue management using neural networks*** . Sharif Azadeh, Sh, R. Labib, and G. Savard. *International Journal of Revenue anagement* 7.1 (2013): 18-36.