An Overview of Intelligent Systems Technologies for Process Optimization

Edward Wilson, Ph.D., P.E.

Ed.Wilson@intellization.com

Intellization™
INTELLIGENT SYSTEMS FOR OPTIMIZATION

presented at "Evolving Computer Application Trends in the Steel Industry" on May 8, 2000 in Myrtle Beach, SC. Symposium sponsored by the Iron and Steel Society's Computer Applications Committee

Situation:

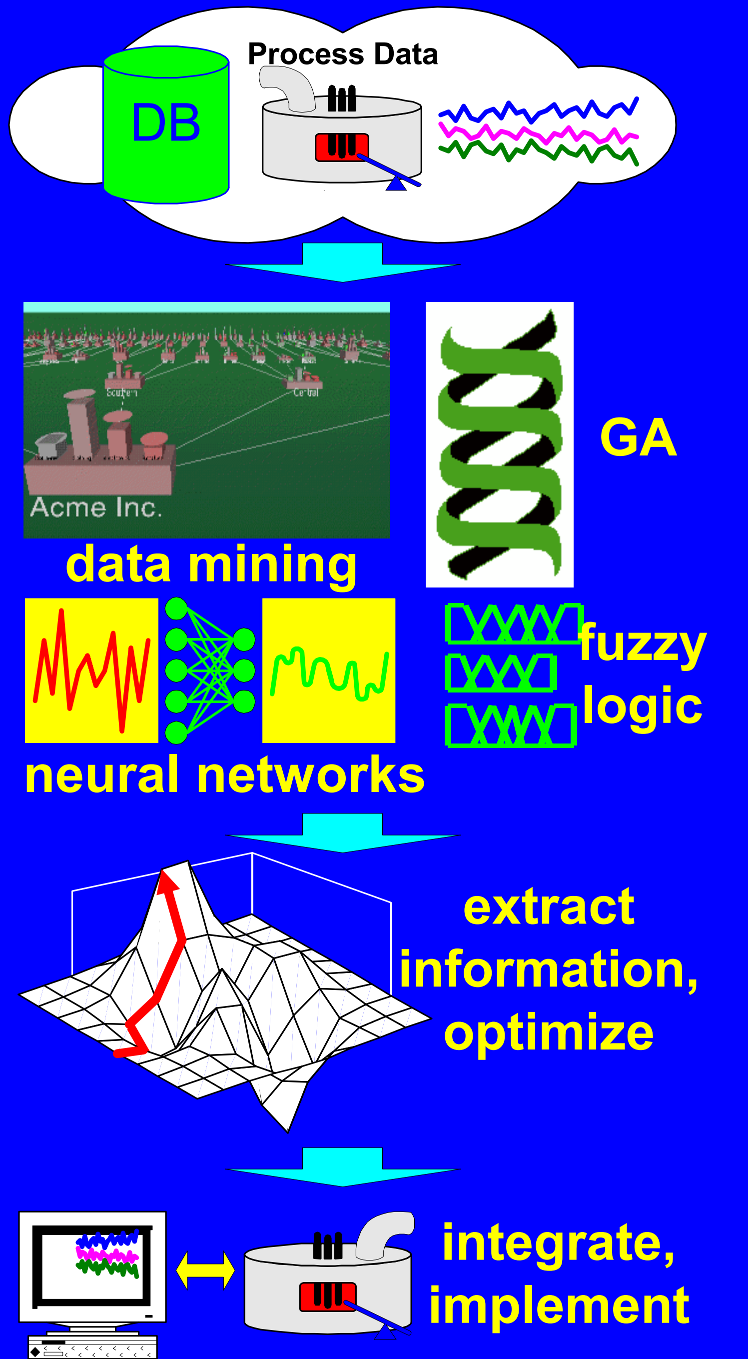
- You've got a **database** full of historical information about your plant.
- You've got many **sensors** giving real-time information about your process.
- You'd like to cut costs, increase efficiency, make better decisions.
- You know there's useful **information** in the **data** that could help run the process better, but don't know how to get it.

Goal:

- **Get that information** to help you make **better decisions** or to **optimize** your control system.

Solution:

- Develop a system that integrates **intelligent systems technologies** with the control and information management systems to **optimize your process**.



Intellization

Goal of Presentation

- Give you information about these intelligent systems technologies that will help you understand them and see how you might use them to improve your process.
  - Overview of specific technologies
  - Issues common to intelligent systems in general
- How to choose the right tool for the job.

Intellization

Outline

- Intelligent systems (IS) introduction
- Technology overview
  - Data mining
  - Neural networks
  - Genetic search
  - Fuzzy logic
- General IS application issues
- Application examples
- Application development guidelines
- For further study …

Artificial Intelligence

- One definition: “the attempt to build machines that think and act like humans, that are able to learn and to use their knowledge to solve problems on their own”
- Another one: “a computer program that produces results that could not be *directly* programmed by its creator”

- My bias is towards practical near-term value vs. science fiction: How can AI be used to optimize performance and save money? …
- Different maturity levels among various methods

Intellization

Information from Data

- Gauss – mathematics to fit astronomical observations to determine orbits of planets
- … (a couple of hundred years go by)
- Today – developing algorithms and technology, using computers to **increase the level of automation** in extracting information from data

- Converging technologies enable these advances:

1. **Increasing computing power**
2. Theoretical, statistical, learning algorithm advances
3. Advances in data collection, management, and communication (networking, internet)

(These all complement one another)

GPI's 1000-node Beowulf Cluster

- •Began operation in July 99
- •1000 350 MHz Pentium IIs
- •500 Dual processor motherboards
- •Individual cases
- •100 Mbit/s Ethernet
- •25 20-port hubs
- •2 network switches
- •Linux OS
- •Beowulf architecture

- Genetic Programming, Inc. (Los Altos Hills, CA) built this and uses the parallel computing power to **evolve** solutions to optimization problems (discussed later).
- Example of how computing power enables a **completely new approach** to solving problems.

Intellization

Outline

- Intelligent systems (IS) introduction
- **Technology overview**
  - **Data mining**
  - **Neural networks**
  - **Genetic search**
  - **Fuzzy logic**
- General IS application issues
- Application examples
- Application development guidelines
- For further study …

Technology Overview

- Data mining
  - Extracting useful information from data warehouses. Based on statistics.
- Neural networks
  - Signal processing, optimization method, for static or real-time data
- Genetic search
  - Optimization method based on survival of the fittest.
- Fuzzy logic
  - Implementation method, static or real-time

Data Mining

Definition: The **automated** extraction of hidden, previously unknown, predictive information from databases (and then using it)

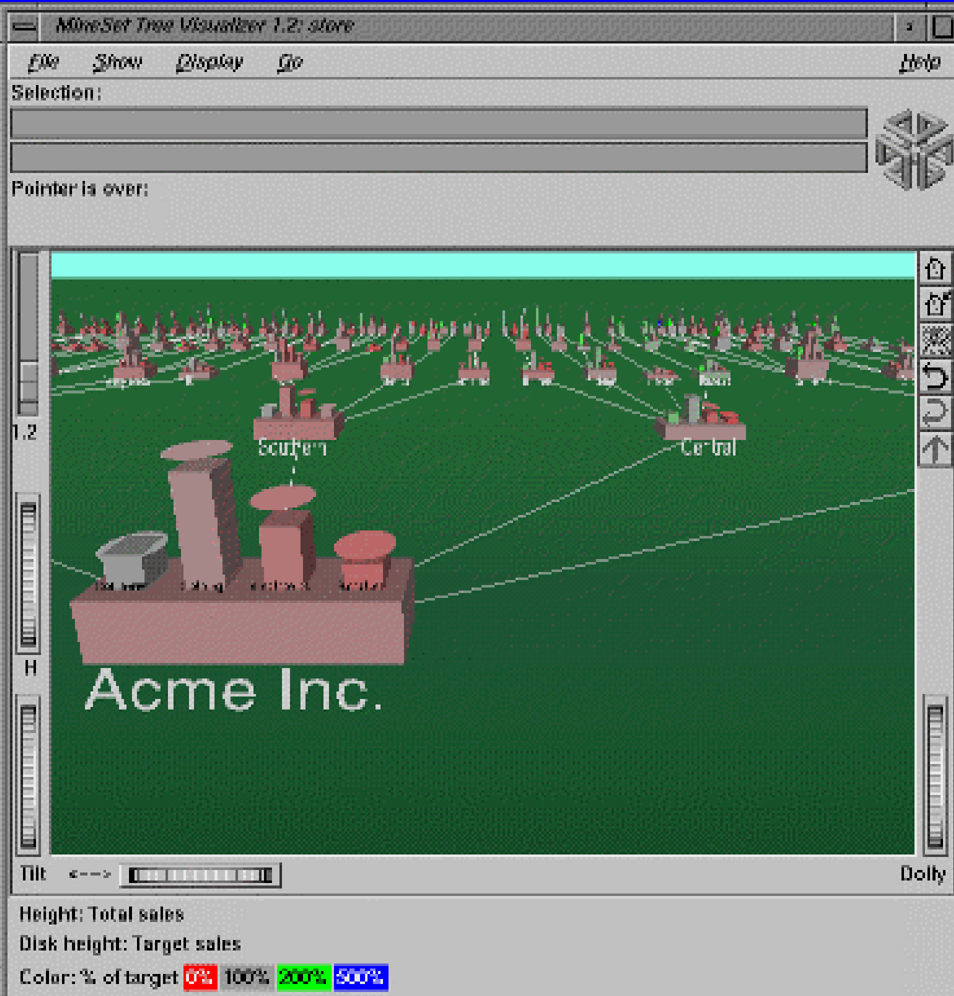
- A key issue here is the **degree** of automation. Goal of automation is to help/replace the statistician, and to some degree, the domain expert.

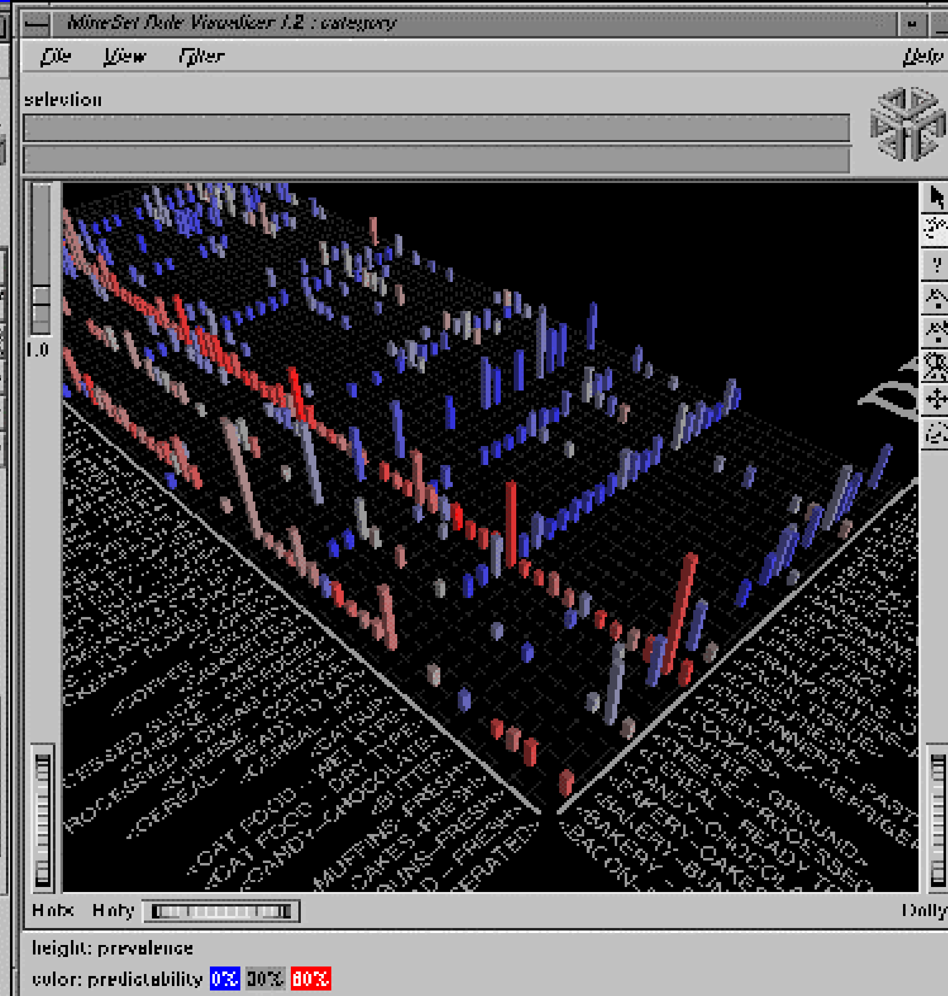
- Data mining technologies:
  - Decision trees, rule induction, k-nearest neighbor algorithms, neural networks, genetic algorithms
- Related to, but not considered DM:
  - Data warehousing, reporting, OLAP, data visualization

Data Mining Buzzwords

- KDD – knowledge discovery in databases
- OLAP – on-line analytical processing
  - Show me the kWh/ton for my company's small furnaces last month.
  - Architecture optimizes efficiency for processing multi-dimensional data. OL"A"P vs. OL"T"P - optimized for analysis vs. transaction.
- Decision trees, rule induction – automatically generated tree structure or rules used for classification
  - If casting rate > 60 ipm and water pressure below 90 psi, then billet defect #2 is likely
- k-Nearest Neighbor – memory-based prediction / classification
- Data visualization – aid to manual data mining – next slide

SGI Visualization Tools





Visualization enables user to manually explore the data (instead of trying to fully automate the knowledge discovery process)

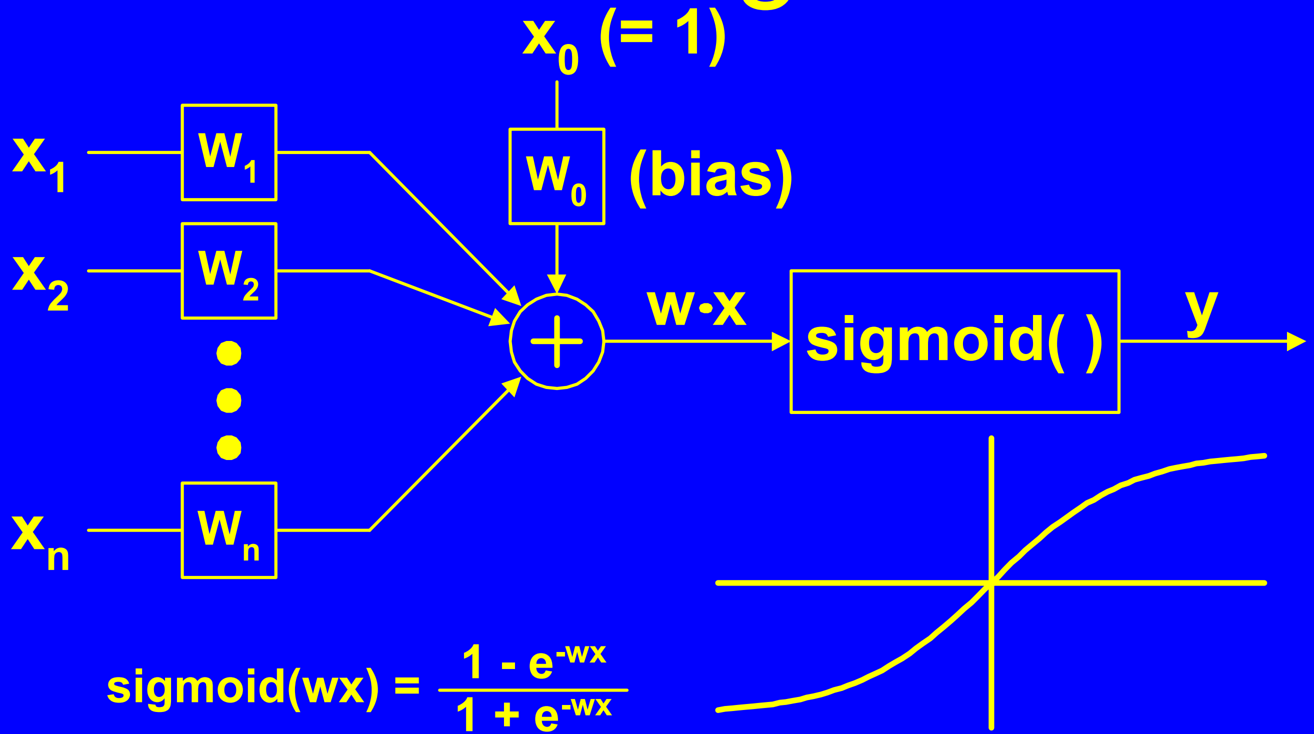
Neural Networks

- Neural Networks (NNs) are signal processing elements, loosely modeled after the human brain, that have valuable capabilities such as:
  - Nonlinear → deal with real-world
  - Adaptive → **"trained" with data** to solve problems, adapt to changing systems
  - Parallel architecture → fast in hardware
  - Generic functional element → can model anything
- However …
  - These benefits must be weighed vs. costs: black box, nonlinear optimization required

Biological Motivation, Engineering Application

- Human brain:
  - Massively parallel network of simple processors, with great capabilities
  - 15 billion neurons
  - 10,000 inputs per neuron
  - 1-2 ms neuron response time
- ANNs studied in a variety of fields
  - Engineering
  - Psychology

Model of Single Neuron



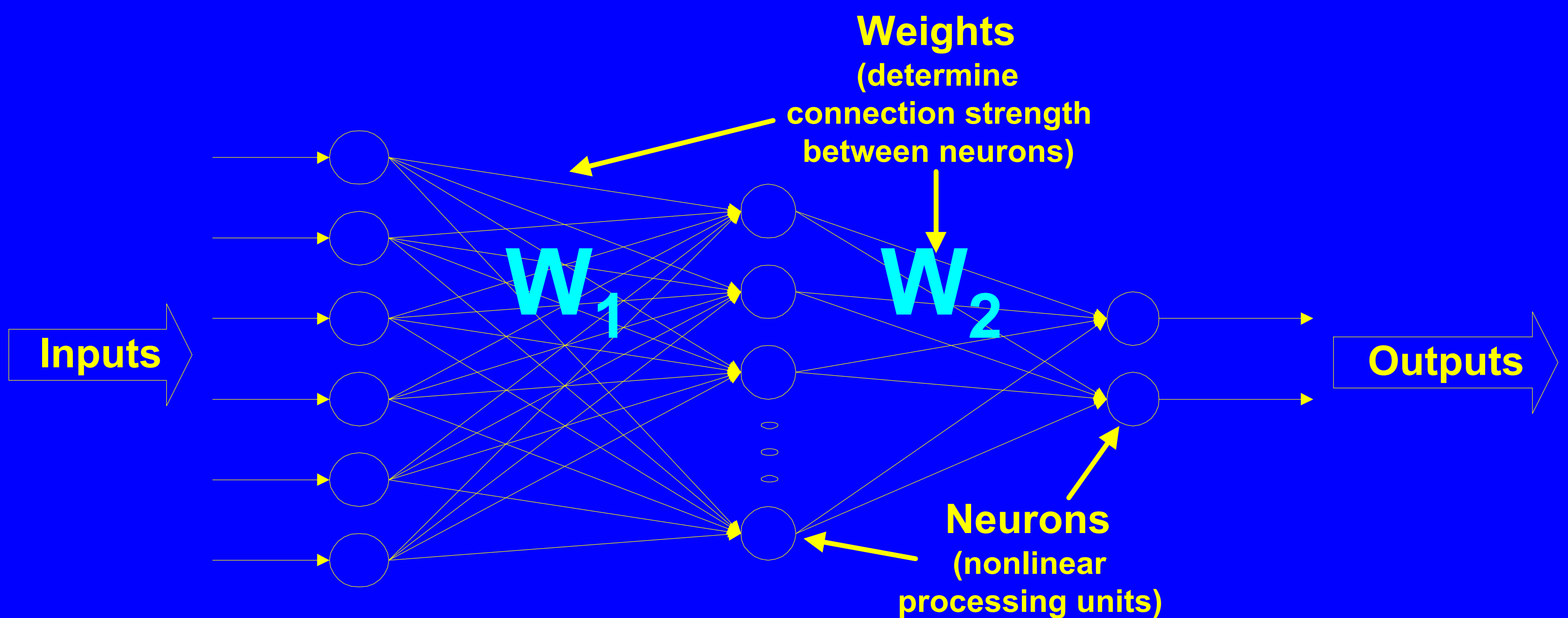
$$\text{sigmoid}(wx) = \frac{1 - e^{-wx}}{1 + e^{-wx}}$$

$$y = \text{sigmoid}(w_1x + w_2x + \ldots + w_nx + \text{bias})$$

- Implement in software or hardware
- Loosely modeled from biology, but chosen for processing and training
- This type most common for engineering applications

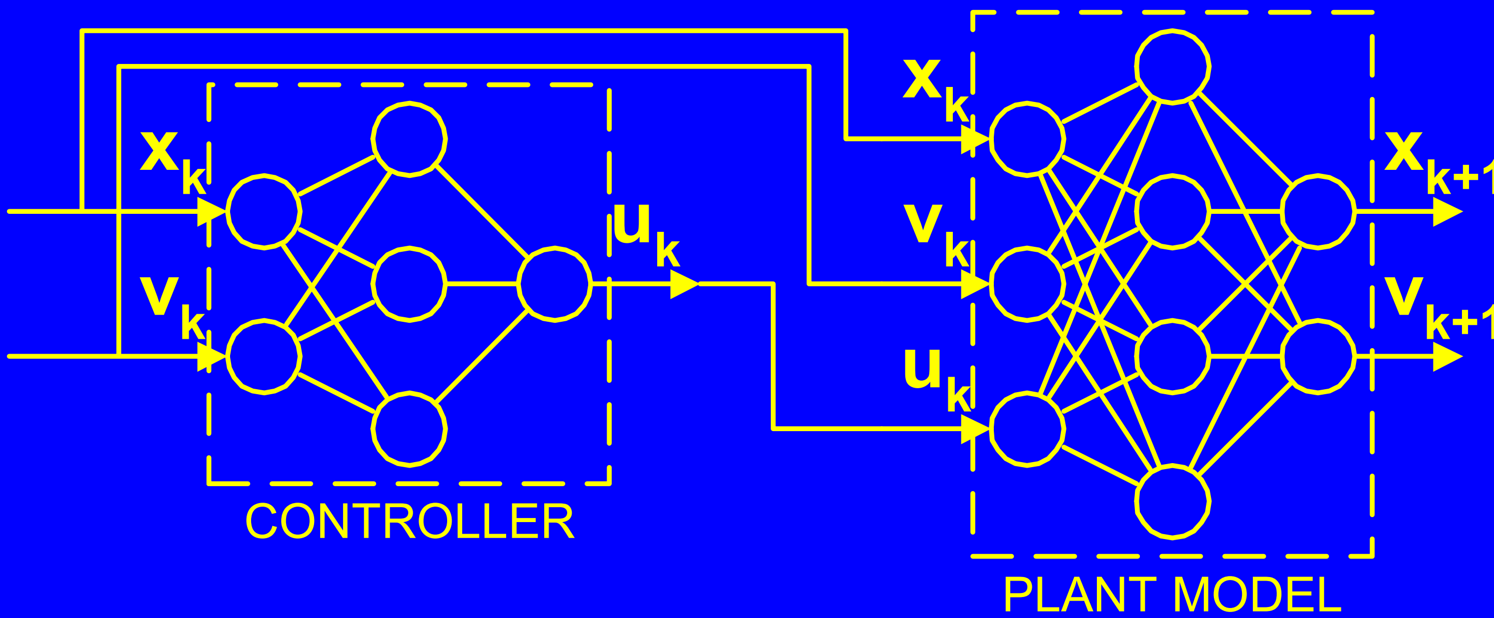
Intellization

Model of Neural Network



- output = W_2*sigmoid(W_1*inputs)
- Proven to be a generic nonlinear functional element
- Functionality defined by architecture, weights, (training)

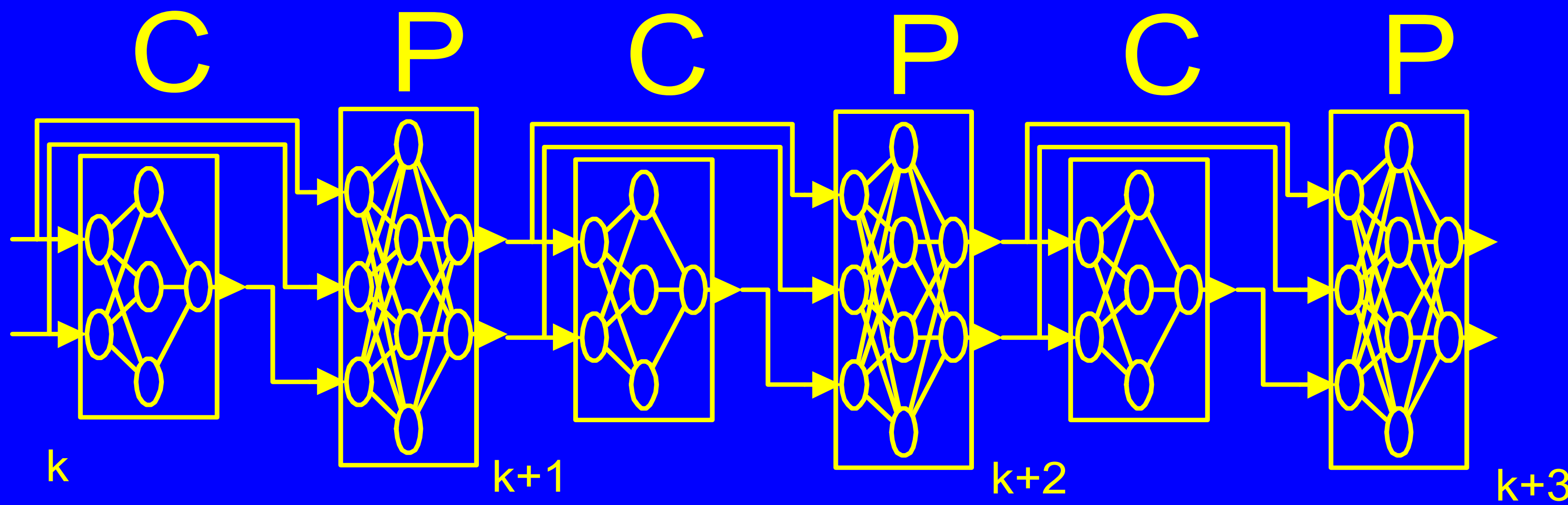
Indirect Training - Controller Optimization



- Plant model is fixed (adapted separately to model plant data)
- Calculate cost function, $f(u_k, x_{k+1}, v_{k+1})$
- BP to get derivatives, optimize weights in controller
- Requires direct, continuously differentiable path from control parameters to cost calculation

Intellization

Controller Opt. - LQR, terminal



- "Backpropagation through time" - copies controller (C), plant (P)
- Cost on terminal state, intermediate states, actuators, etc.
- BP to find derivatives of cost w.r.t. controller parameters. Then optimize.

Neural Networks - Summary

- Generic nonlinear functional element - can implement any MIMO mapping function to arbitrary accuracy (universal approximator)

+ “Trained” with data – uses gradient-based parameter optimization

+ Solid mathematical foundation - BP gets derivatives, then standard gradient-based optimization problem

+ Parallel architecture, but usually implemented in software on serial computer

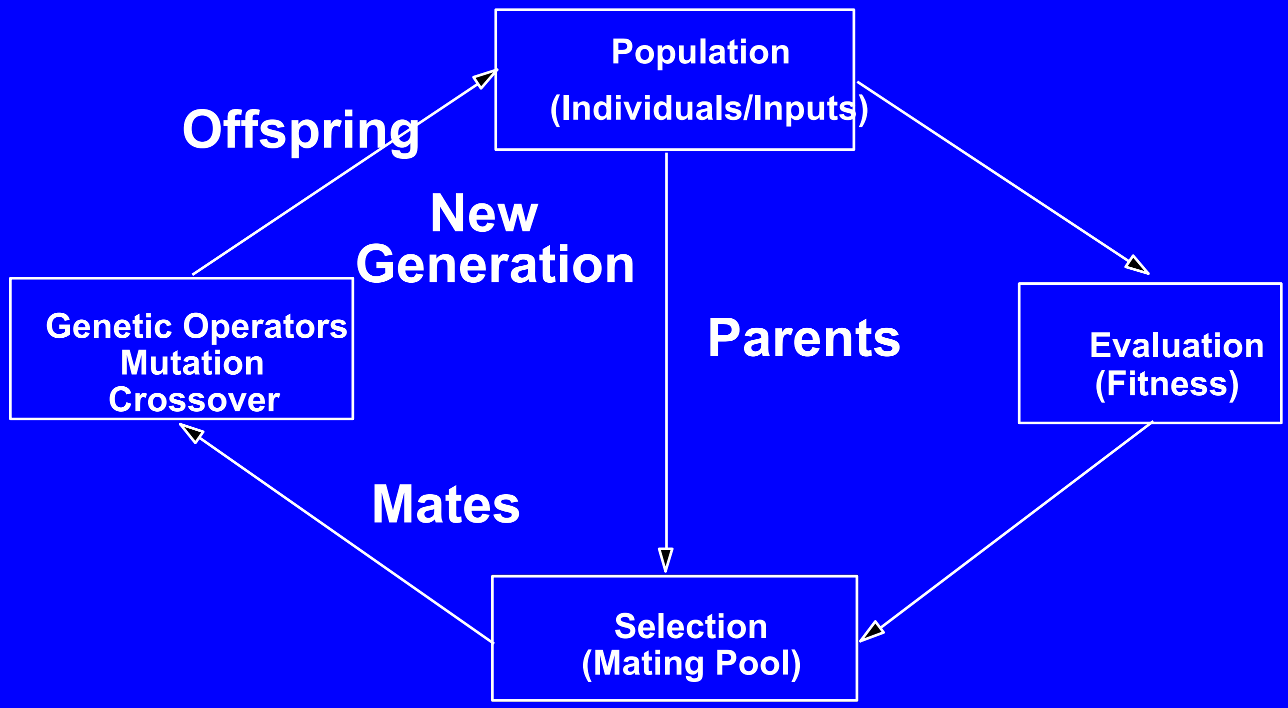
– Black box - difficult to understand inner workings

Genetic Search

Genetic Algorithms (GA), Genetic Programming (GP)

- Uses Darwinian principles of random mutation and crossover to improve solution to a problem.
- The solution (for example, schedule) is encoded in a chromosome representation.
- A population of solutions combine (crossover) and change randomly (mutate) during each generation of offspring.
- Most are discarded but the few that come close to doing the job combine with other survivors and spawn offspring programs-a process that may produce results superior to anything crafted by humans.
- GA: string representation, GP: tree representation
- Designing the encoding (GA vs. GP, then more specifically) is very important.

Genetic Search



\+ Parallel solution possible

\+ Does not require differentiable system model

- Scheduling is most successful application to date
- Traveling Salesman Problem

\+ Not susceptible to local minima in the optimization cost surface

\- Very computationally intensive

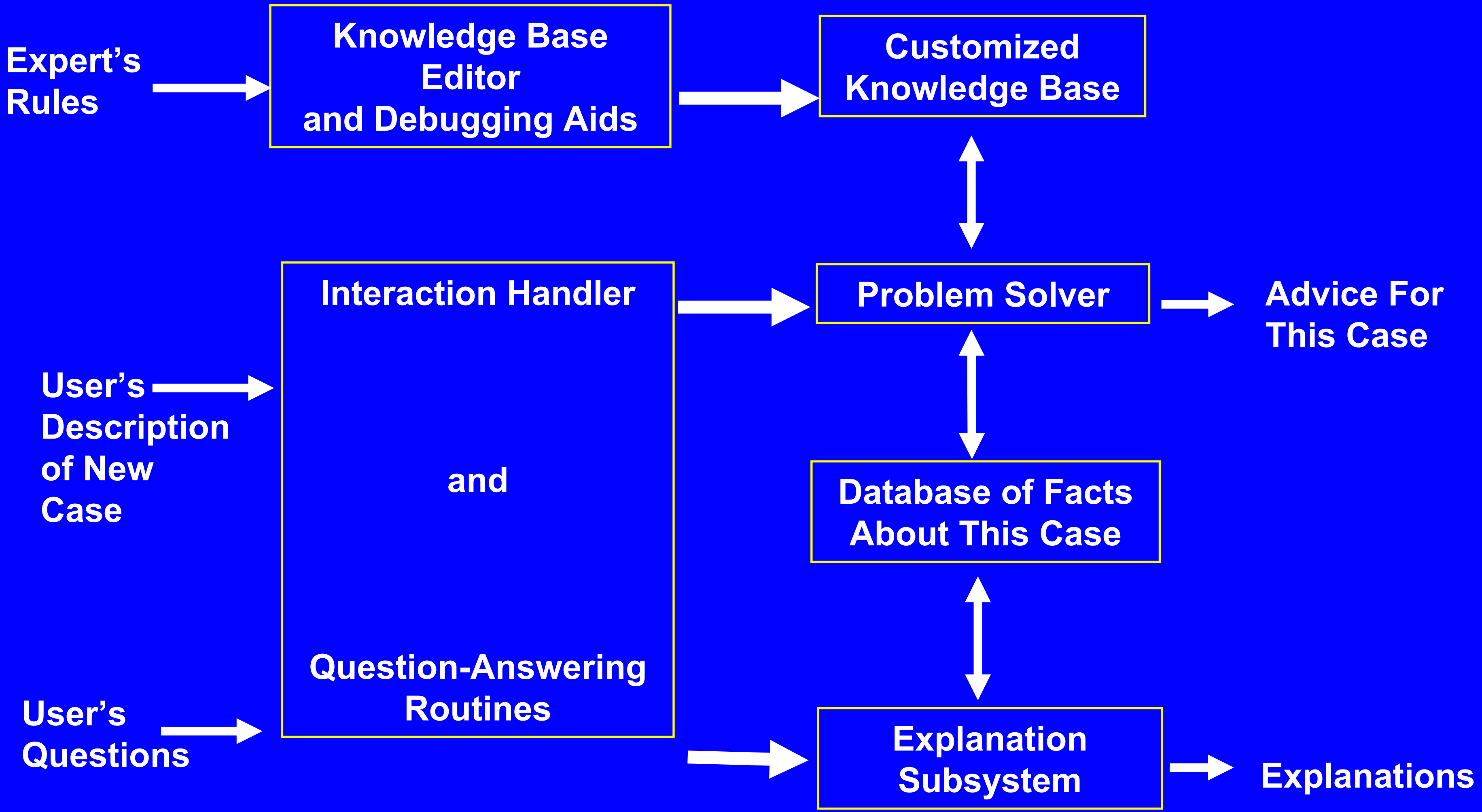
Intellization

Expert Systems

- Expert knowledge of system required
- Hard to translate knowledge to precise rules
- Can be useful if rules are known
- (above can be automated to some extent)

- Medical diagnosis expert systems: MYCIN, INTERNIST
- Industrial Process expert systems: Fault detection and preventive maintenance in the chemical industry (distillation and cracking) by GENSYM

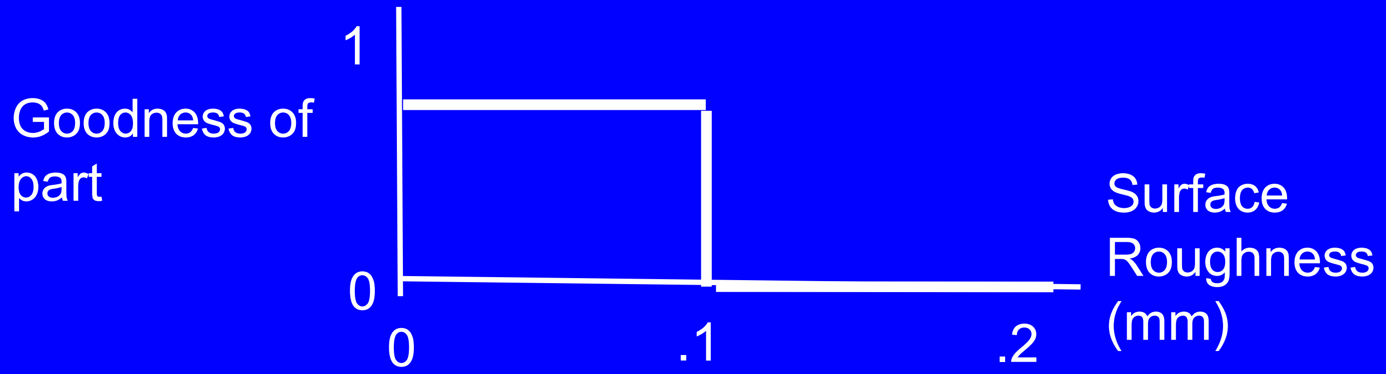
Expert Systems



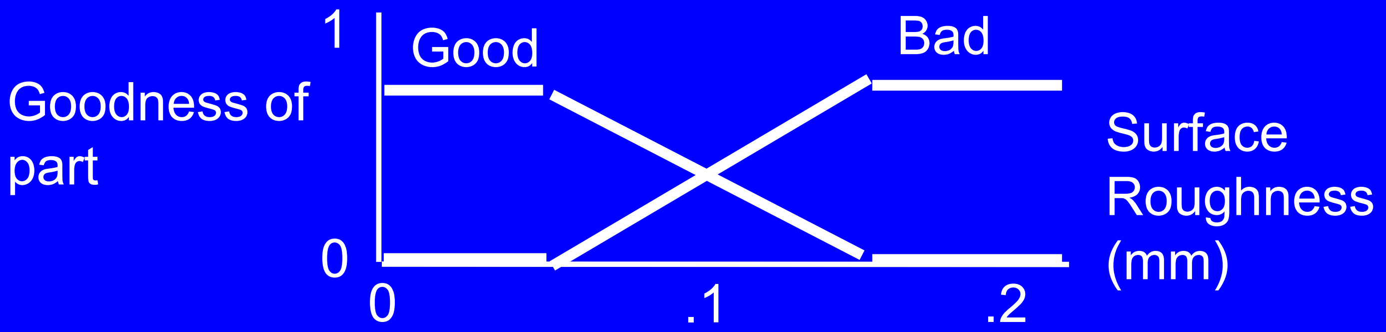
Intellization

Expert System vs. Fuzzy Logic

Expert System



Fuzzy Logic

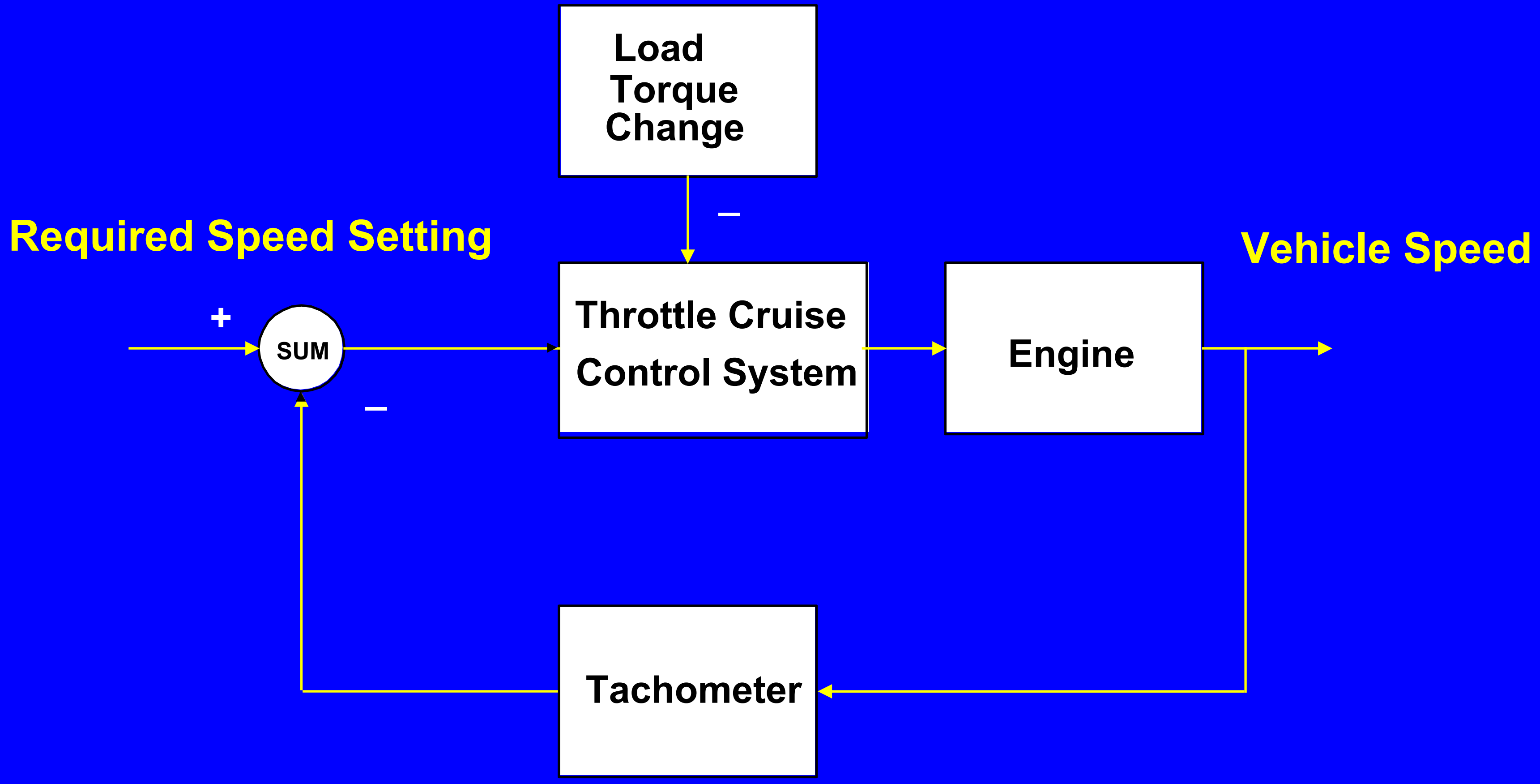


Intellization

Fuzzy Logic

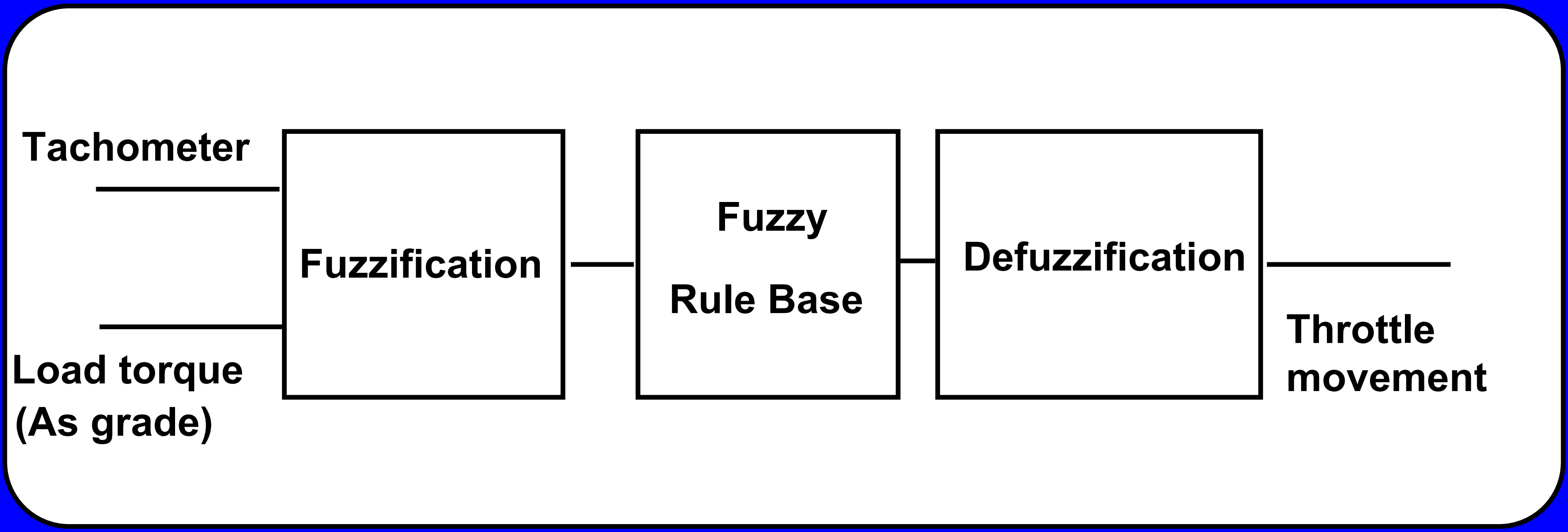
- Rule Based System--Uses linguistic terms
  - light
  - moderate
  - heavy
- Absolute knowledge of system not required
- Responses to a statement reflect the degree to which the statement is true

Fuzzy Logic Cruise Control



Fuzzy Logic

FUZZY THROTTLE CRUISE CONTROL SYSTEM

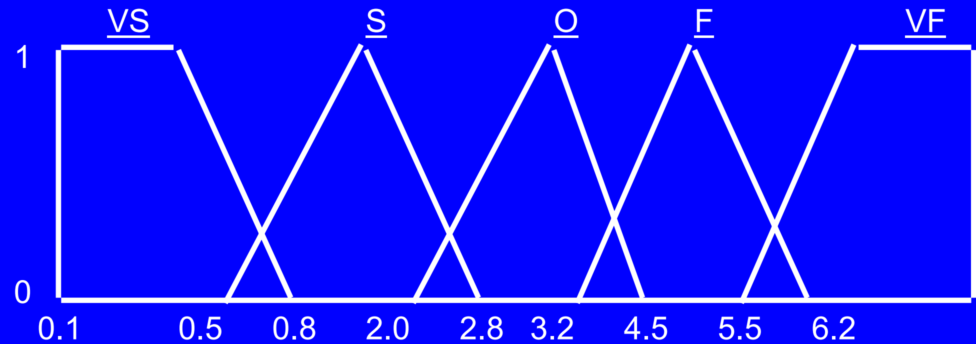


Fuzzy Logic - Set Membership

Tachometer

Rotational speed (rad/sec)

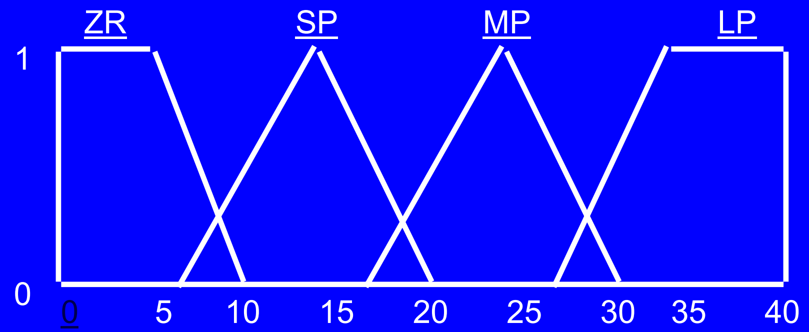
VS = Very Slow
S = Slow
O = Optimal
F= Fast
VF = Very Fast



Load Torque (as grade)

Degree of grade (%)

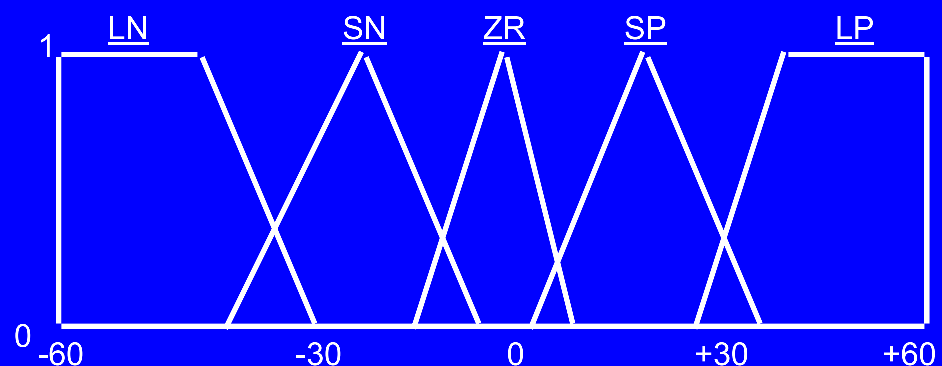
ZR = Zero
SP = Small Positive
MP = Moderate Positive
LP = Large Positive



Throttle Movement

Rotational speed (rot unit/sec)

LN = Large Negative
SN = Small Negative
ZR = Zero
SP = Small Positive
LP = Large Positive

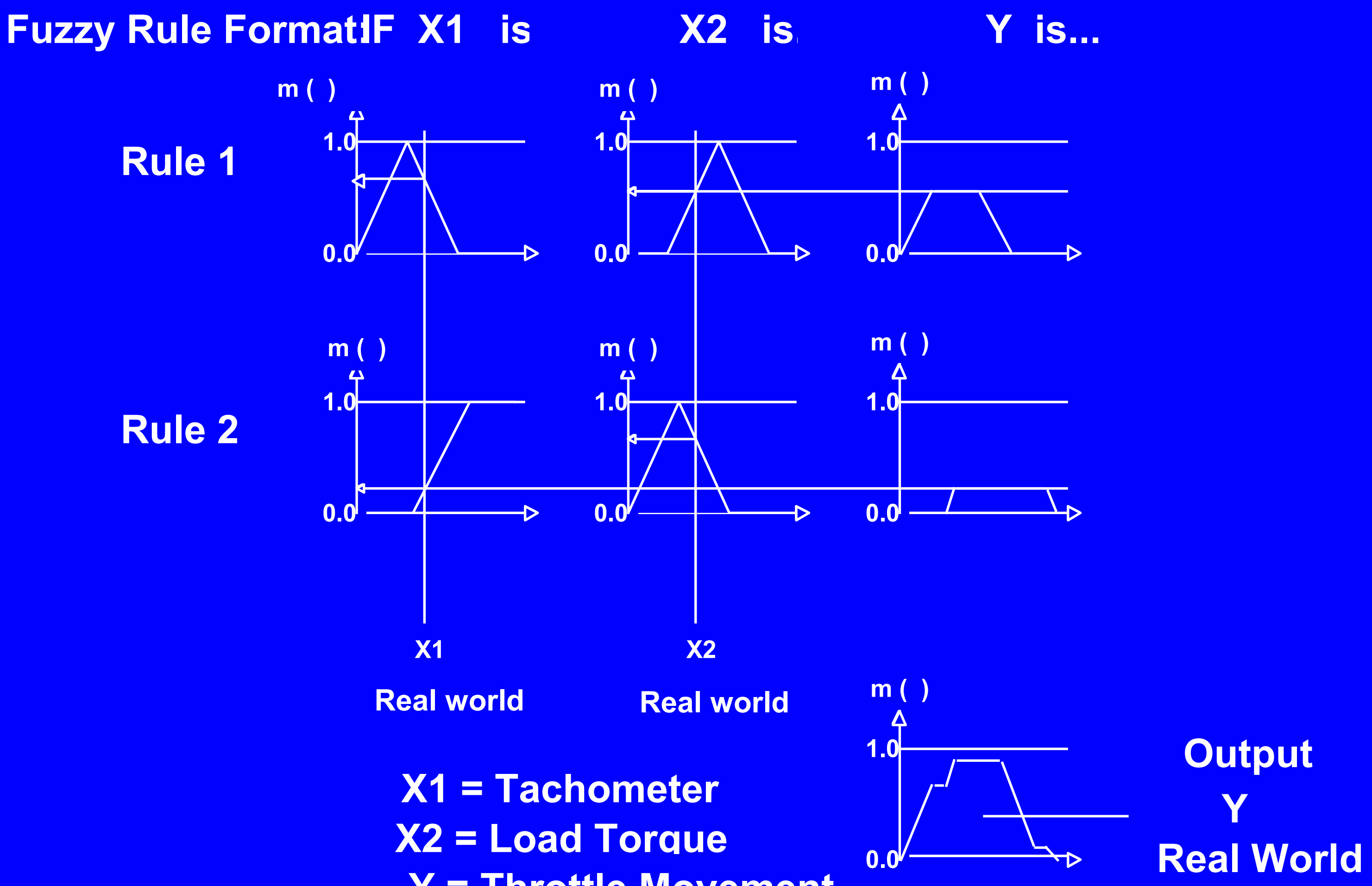


Intellization

Fuzzy Logic

Load Torque \ Tachometer	Very Slow	Slow	Optimal	Fast	Very Fast
Zero	Large Positive	Small Positive	Zero	Small Negative	Large Negative
Small Positive	Large Positive	Small Positive	Zero	Zero	Small Negative
Moderate Positive	Large Positive	Small Positive	Small Positive	Small Positive	Zero
Large Positive	Large Positive	Large Positive	Large Positive	Small Positive	Small Positive

Fuzzy Logic: Inferencing



X1 = Tachometer
X2 = Load Torque
Y = Throttle Movement

Fuzzy Logic Summary

- Useful for implementing control systems
- “Fuzzy” enables smooth blending of outputs, unlike “crisp logic” systems
- Largely manually programmed, although adaptation possible.
- Best used on weakly nonlinear systems where good control is more art than science and there’s an expert operator to learn from (explicitly or adaptively)
- Consistent, as if best operator always working at his best

- Note – this is not as data/compute intensive as the other intelligent systems technologies

Outline

- Intelligent systems (IS) introduction
- Technology overview
  – Data mining
  – Neural networks
  – Genetic search
  – Fuzzy logic
- **General IS application issues**
- Application examples
- Application development guidelines
- For further study …

Issues Common to Intelligent Systems Applications

- Ensuring data integrity
- Building/using the model, data mining engine, etc.
- Pre-filtering data
- Selecting model order

The same technology convergence (compute power, …) discussed earlier helps address these issues.

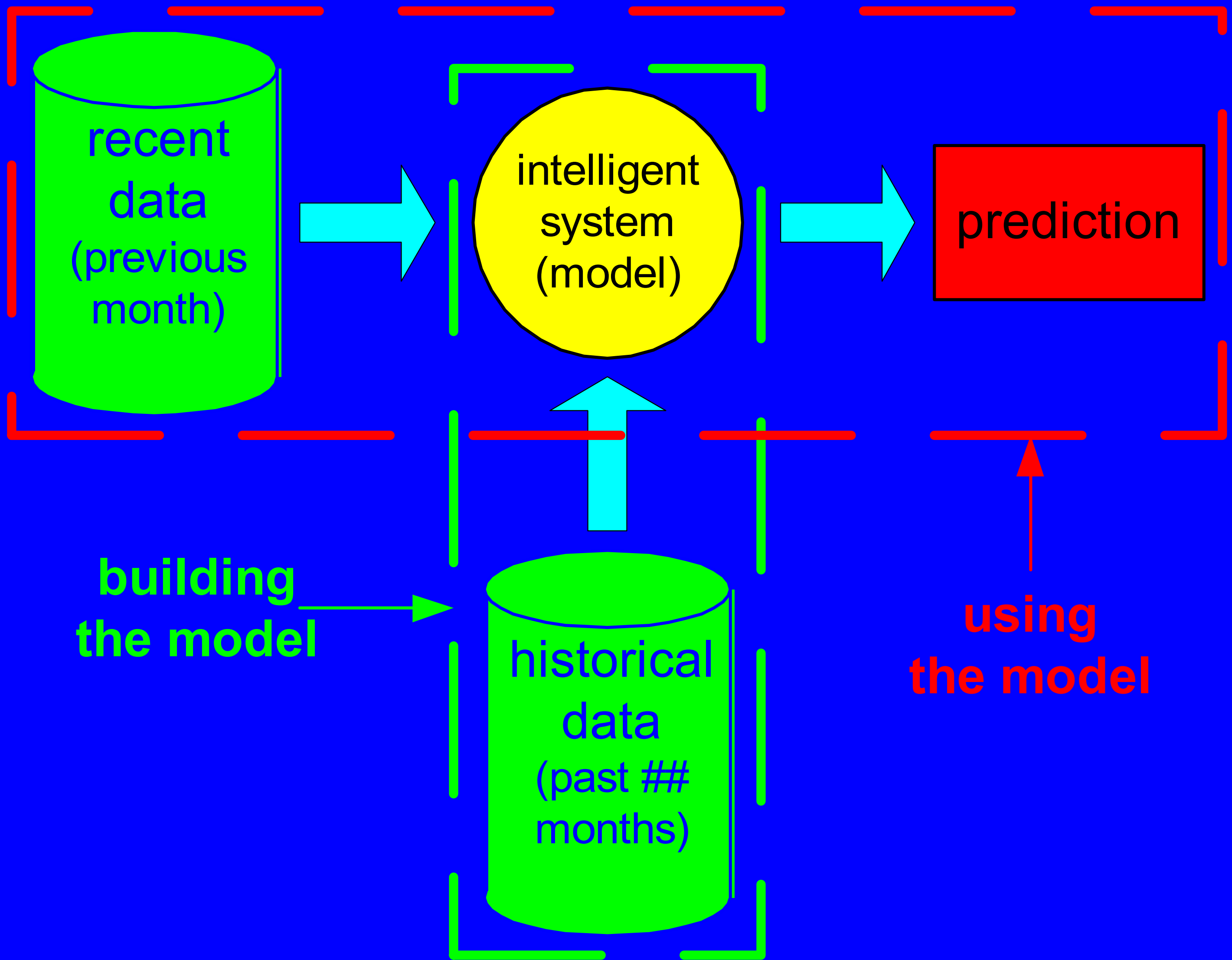
Data Collection, Integrity …

The IS may be critical to providing a competitive advantage, but it is just one part of the overall solution. This is why successful IS products tend to be application-specific (vertical vs. horizontal – like Excel or SuperSolver discussed yesterday)

- Sensors
- Experiment design (when needed)
- Data collection
- Ensuring data integrity (GIGO)
- Filtering, pre-processing
- **Intelligent system development**
- Integration
- Usability

Increased automation ➔ increased dependence on data (vs. myth that intelligent systems are immune to bad data)

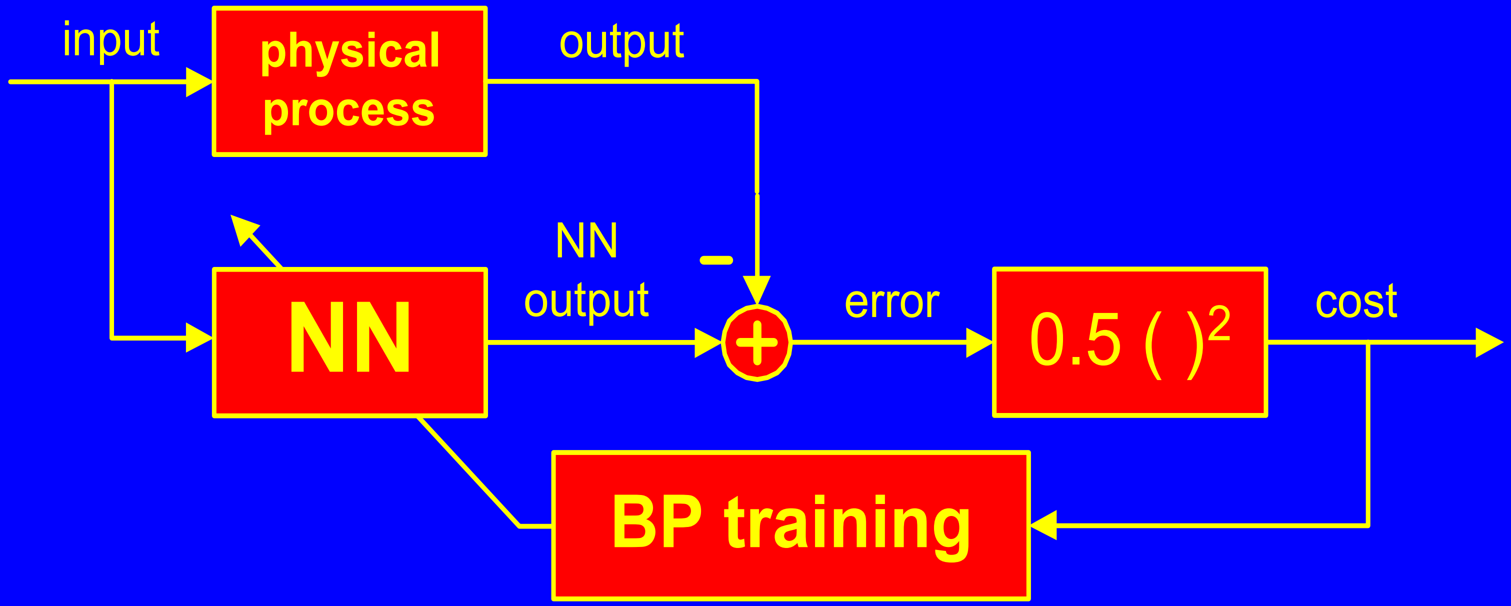
Building and Using a Predictive Model



Intellization

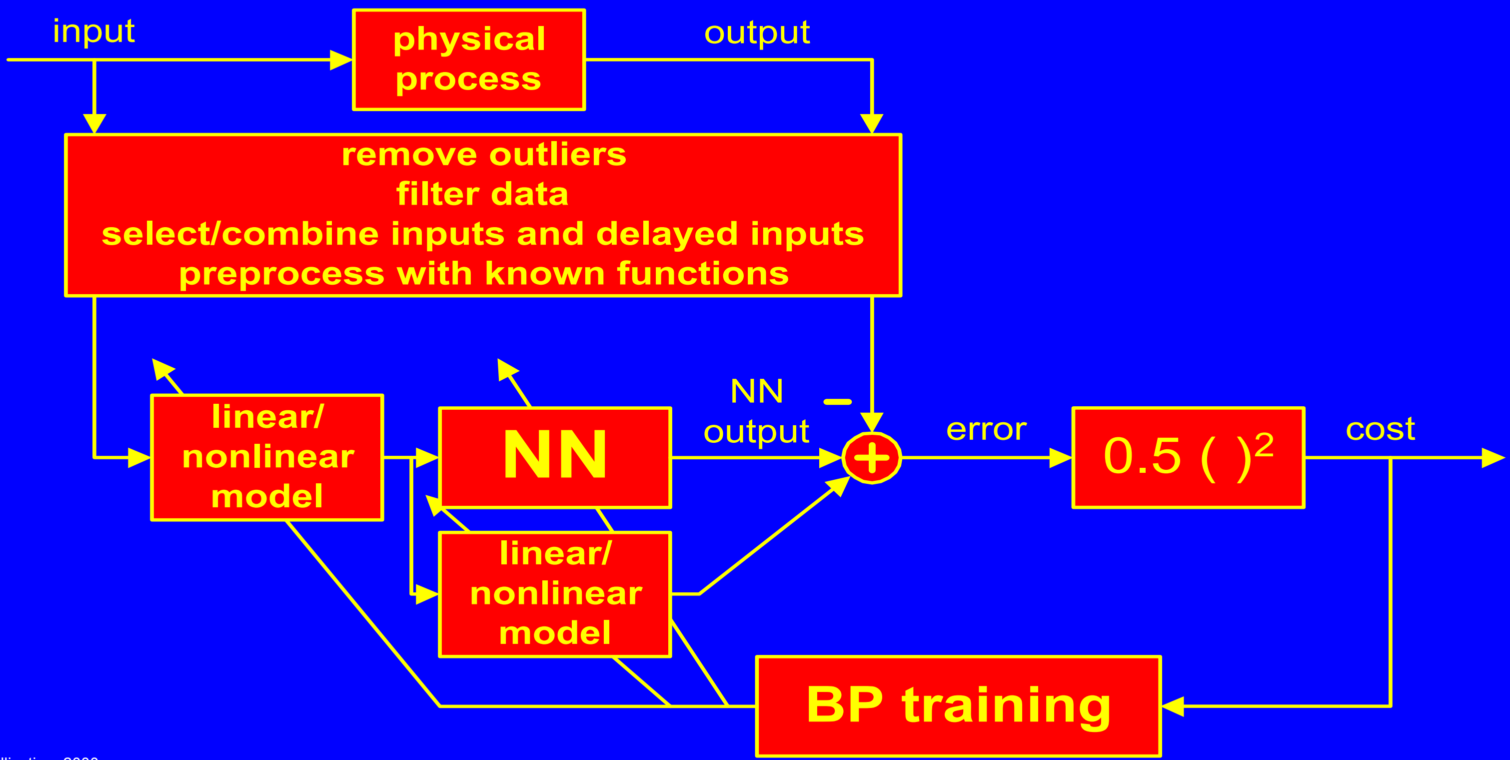
Data Pre-processing Example: Training NN Model

- NN trained to emulate a physical process
- Once you have this predictive model, it can be used in control, etc.



Data Pre-processing, Structure

- System ID issues: selection of inputs, sample rate, sufficient persistently exciting data, sufficient dof in model, remove outliers, etc.
- More complex than it first seemed. Understanding of underlying process physics is important in designing the pre-processing and the overall architecture. This is a relatively significant effort.



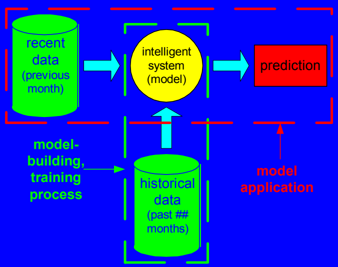
Intellization

Model Order Selection

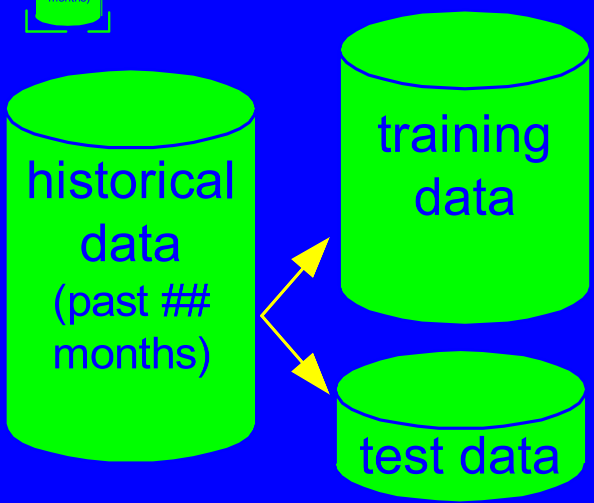
- Occam's razor – "The simplest theory that fits the facts of a problem is the one that should be selected." - William of Occam (1284-1347)
- "Things should be made as simple as possible, but not simpler." - Albert Einstein

- Especially important for NNs (and other intelligent technologies) – many degrees of freedom, noisy data
- Higher order will always produce a better fit, but this may not be the better model – "over-fitting"

Order Selection (and IS development)



- There's a need for testing IS performance during development, and also before going on-line



1. use this data set to develop the prediction algorithm (including model order, inputs, etc.)

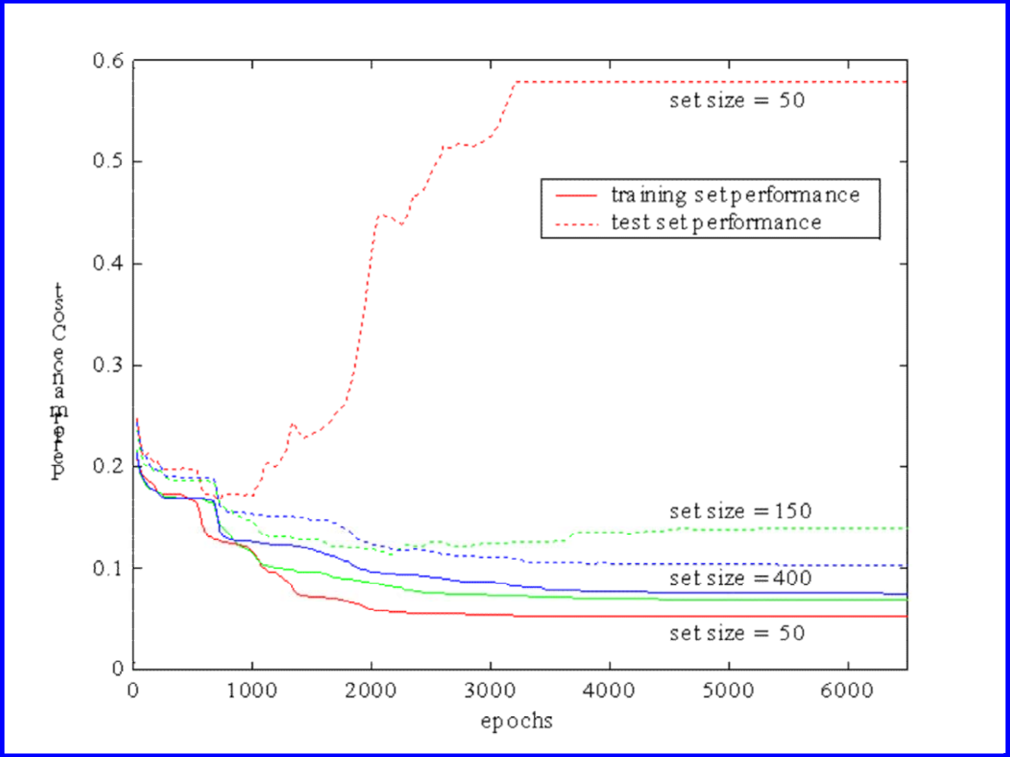
2. use this data set to test the resulting algorithm

3. use the test results to improve algorithm design

- If model contains too many dof, the IS will adapt to noise rather than the true underlying physics. But the fit will be very good.
- When testing on new data, fit will be bad. Performance on test data drives IS development. Not foolproof, however.

Intellization

Over-fitting Example



- Horizontal axis – training iterations
- Vertical axis – neural-network performance, lower is better
- Training performance improves monotonically
- Test performance (what really matters) worsens once over-fitting begins – e.g., around 1000 iterations for set size = 50
- Over-fitting error can be reduced by using larger data sets

Intellization

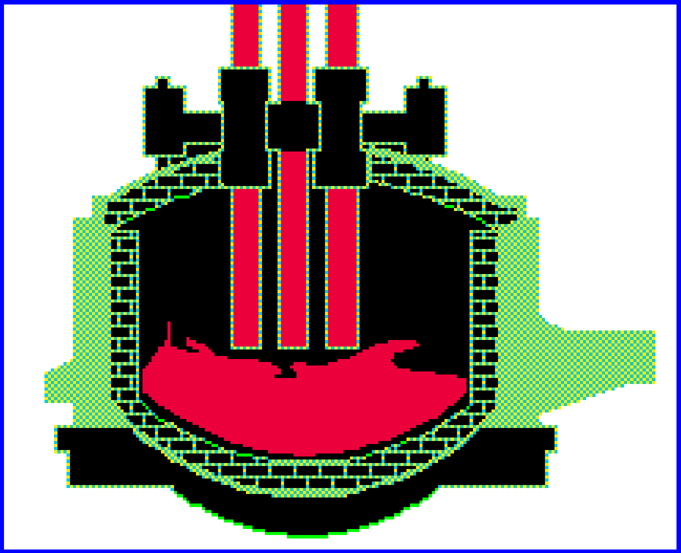
Automating Model Order Selection

- MATLAB demo
- Highly automated generation of a dynamic model for an industrial process, using real data – Hydraulic drive for EAF electrode

Outline

- Intelligent systems (IS) introduction
- Technology overview
  - Data mining
  - Neural networks
  - Genetic search
  - Fuzzy logic
- General IS application issues
- **Application examples**
- Application development guidelines
- For further study …

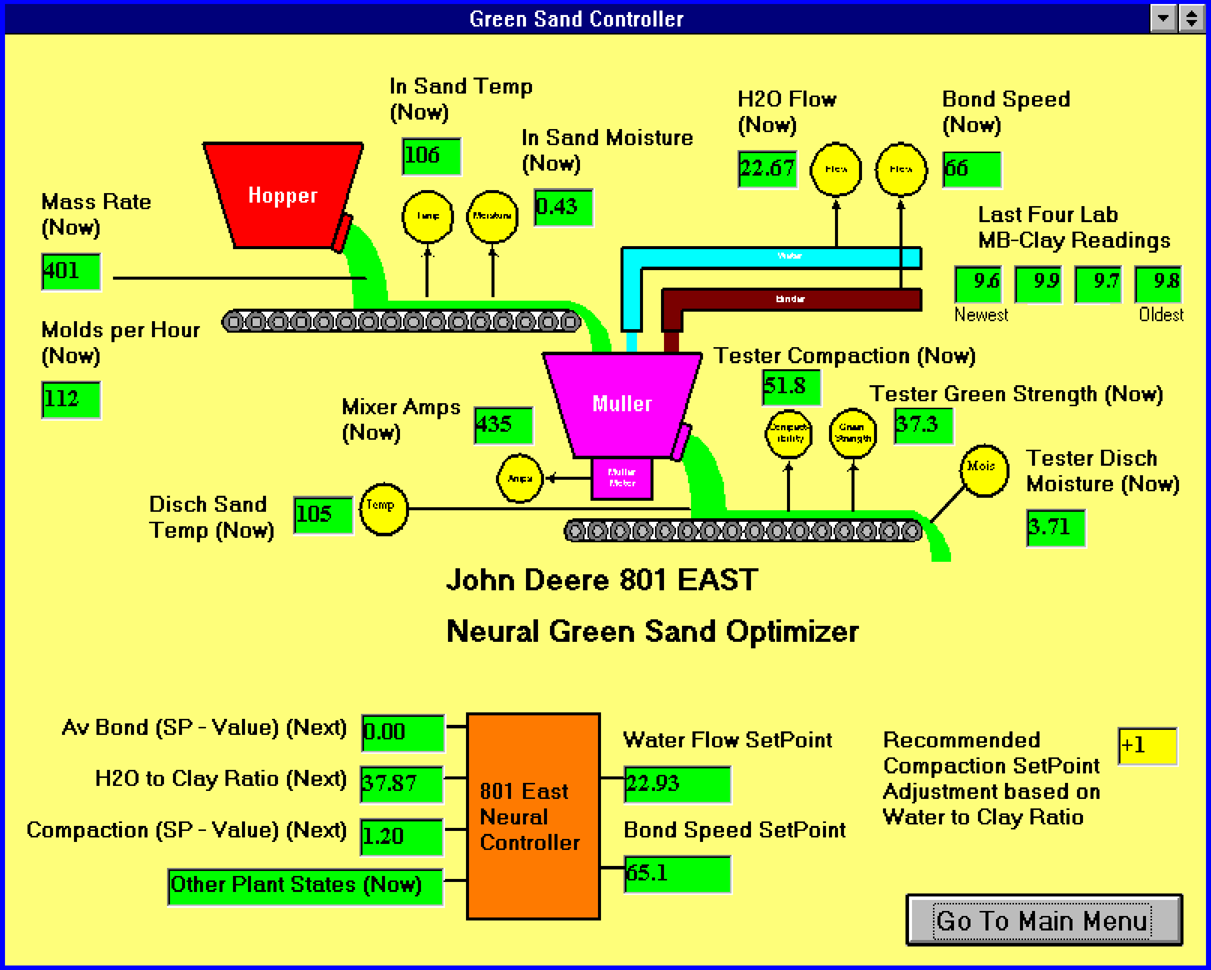
Intelligent Arc Furnace Controller



Reduces production costs:
- less energy
- less electrode consumption
- increased throughput

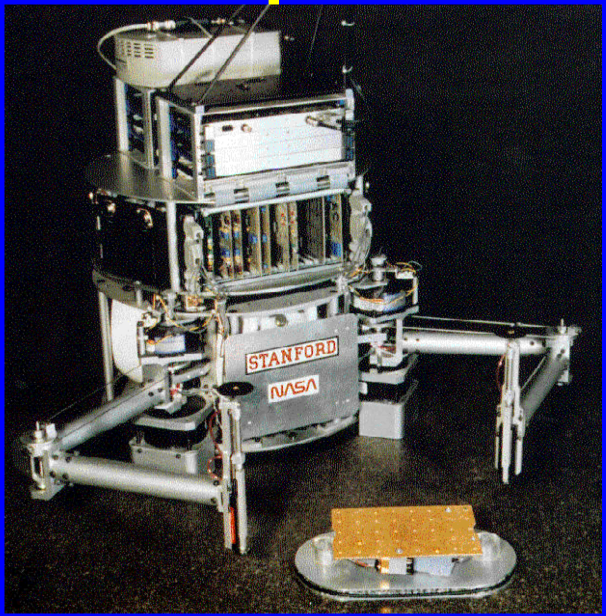
- Originally developed by Neural Applications Corp., now SAI
- NN-based regulation system for control of electric arc furnace, used in steelmaking
- Uses backpropagation, similar to model-based predictive control architecture
- Adapts to changes in charge and system dynamics
- Data availability outweighs process understanding
- 40 installations worldwide - each one is different

Green Sand Controller



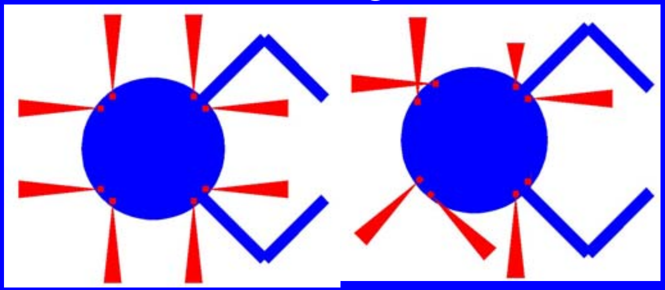
- NN predicts process output to account for transport delay - used as “virtual sensor”
- Neural Applications Corp./SAI application

Neural-network application to spacecraft/robot Control

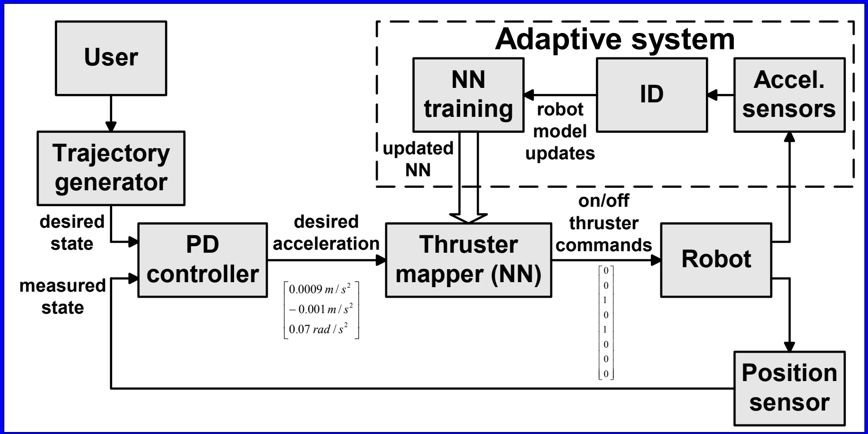


Stanford free-flying space robot

Thruster configurations



Nominal, then after multiple unknown failures (some destabilizing)



Intellization

Outline

- Intelligent systems (IS) introduction
- Technology overview
  - Data mining
  - Neural networks
  - Genetic search
  - Fuzzy logic
- General IS application issues
- Application examples
- **Application development guidelines**
- For further study …

Intelligent System Pre-Development

- Beware the hype. Realistically understand and consider the capabilities of the intelligent technologies.
- Consider your eventual goal before building database or architecting the control system (if possible). Design your data collection to maximize data integrity (e.g., use mission critical sensors since they will be maintained)
- Verify your data integrity, filter, correct if needed
- First, try some “simpler” things:
  - Classical, modern control approaches
  - Manual data mining
  - OLAP
  - Visualization
  - Linear regression
  - Decision trees, statistics to find obvious relationships
- Finally, use NNs or GAs to do what the other methods could not.

Intelligent System Development

- Cost / benefit analysis for the IS technologies
- Evaluate, solve using "conventional" methods.
  - Provides a platform for IS development
  - Points to the parts of the problem that need help (e.g., nonlinearity)
- Structure total solution to use IS in conjunction with "conventional" methods
- Control example – build up from fixed linear controller ➔ adaptive NN
- Hard to determine level of difficulty until you're deep into the data, so plan for varying levels of "modelability"

Which Intelligent Technology?

- **Data mining** (decision trees, rule extraction): Use with a database of numerical and non-numerical data to make classifications and predictions.
- **Neural network**: Use when data availability outweighs process understanding. Prefers raw numerical data. Nonlinear modeling and control of physical processes.
- **Genetic search**: Use when need to optimize something over a large search space containing local minima. Have a model, but is not differentiable.
- **Fuzzy logic**: Use for weakly nonlinear systems to automate the performance of an expert operator who can be interviewed or observed

For further study …

- Data mining:
    http://www.kddnuggets.com
- Neural networks:
    http://ieee-nns.org/
- Genetic programming:
    http://www.genetic-programming.com
- Genetic algorithms:
    http://www.aic.nrl.navy.mil/galist/
- Fuzzy Logic:
    http://www.austinlinks.com/Fuzzy/
- Applications:
    http://intellization.com/files/
    Ed.Wilson@intellization.com

Acknowledgment: Many Genetic Algorithms, Expert Systems, and Fuzzy Logic slides were originally developed by Dr. Santosh Ananthraman, Neural Applications Corp.

Summary

- Intelligent systems can be useful for extracting information from data – used to optimize processes or decisions
- Convergence of computing technologies is enabling great advances in this area
- Many issues to consider when working with data
- Variety of technologies – should be chosen carefully and integrated with conventional approaches