2040 Visions of Process Systems Engineering

A Symposium on the Occasion of George Stephanopoulos's 70th Birthday and Retirement from MIT

June 1-2, 2017

Real-time Decision-making under Uncertainty in 2040

Stratos Pistikopoulos

TEXAS A&M ENERGY INSTITUTE
TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE
TEXAS A&M UNIVERSITY

2040 Visions of Process Systems Engineering

A Symposium on the Occasion of George Stephanopoulos's 70th Birthday and Retirement from MIT

June 1-2, 2017

Real-time Decision-making under Uncertainty in 2040

An Adventure in Big Data Analytics or an Opportunity for Model-based Optimization?

Stratos Pistikopoulos

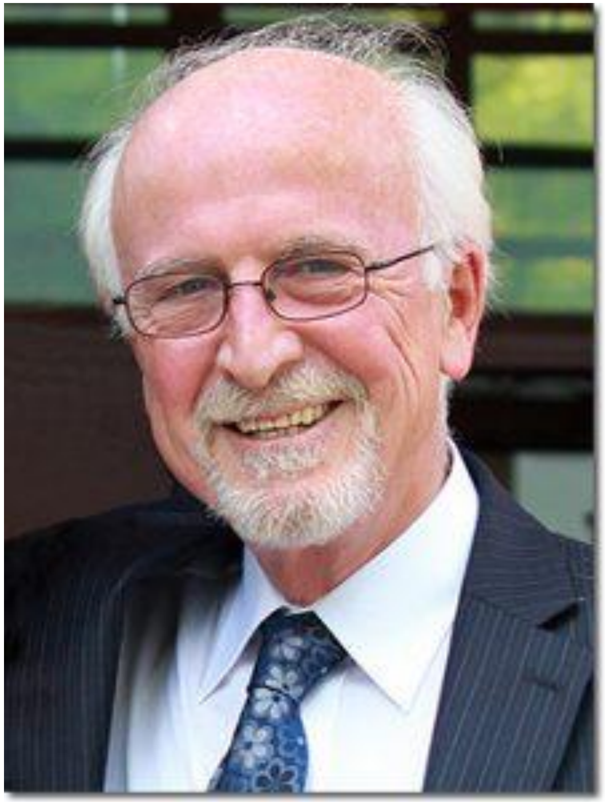
TEXAS A&M ENERGY INSTITUTE
TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE
TEXAS A&M UNIVERSITY

Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Professor George Stephanopoulos

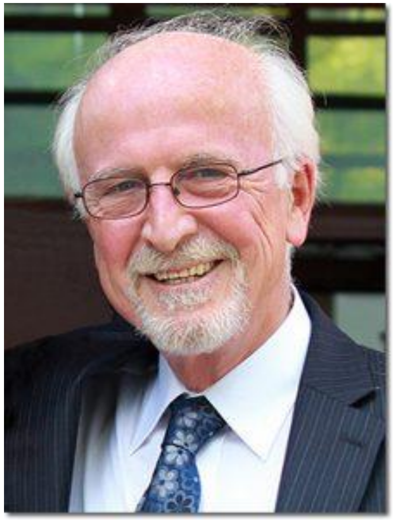


- Research inspiration & impact on my career
- Decision making under uncertainty in 2040

Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Research Inspiration #1



G. Stephanopoulos

Synthesis of process flowsheets: an adventure in heuristic design or a utopia of mathematical programming?

R.S.H. Mah and W. Seider (Editors), Foundations of Computer-Aided Chemical Process Design, 2, Engineering Foundation, New York (1981), p. 439

Artie McFerrin Department of
CHEMICAL ENGINEERING
TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE
TEXAS A&M UNIVERSITY

Research Inspiration #1

Chemical Engineering Science, 1975, Vol. 30, pp. 963–972. Pergamon Press. Printed in Great Britain

STUDIES IN PROCESS SYNTHESIS—I

BRANCH AND BOUND STRATEGY WITH LIST TECHNIQUES FOR THE SYNTHESIS OF SEPARATION SCHEMES

ARTHUR W. WESTERBERG and GEORGE STEPHANOPOULOS†
Department of Chemical Engineering, University of Florida, Gainesville, FL 32611, U.S.A.

(*Received* 24 *June* 1974; *accepted* 4 *February* 1975)

Chemical Engineering Science, 1976, Vol. 31, pp. 195–204. Pergamon Press. Printed in Great Britain

STUDIES IN PROCESS SYNTHESIS—II

EVOLUTIONARY SYNTHESIS OF OPTIMAL PROCESS FLOWSHEETS†

GEORGE STEPHANOPOULOS‡ and ARTHUR W. WESTERBERG
Department of Chemical Engineering, University of Florida 32611, Gainesville, FL 32067, U.S.A.

(*Received* 24 *June* 1974; *accepted* 23 *September* 1975)

A Review of Process Synthesis

NAONORI NISHIDA
Science University of Tokyo
Tokyo, Japan
GEORGE STEPHANOPOULOS
University of Minnesota
Minneapolis, Minnesota 55455
and
A. W. WESTERBERG
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

AIChE Journal (Vol. 27, No. 3)
May, 1981 Page 321

JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 15, No. 3, 1975

The Use of Hestenes' Method of Multipliers to Resolve Dual Gaps in Engineering System Optimization[1]

G. STEPHANOPOULOS[2] AND A. W. WESTERBERG[3]

Studies in the Synthesis of Control Structures for Chemical Processes

Part I: Formulation of the Problem. Process Decomposition and the Classification of the Control Tasks. Analysis of the Optimizing Control Structures.

MANFRED MORARI
YAMAN ARKUN
GEORGE STEPHANOPOULOS
Department of Chemical Engineering and Materials Science
University of Minnesota
Minneapolis, Minn. 55455

AIChE Journal (Vol. 26, No. 2)
Page 220 March, 1980

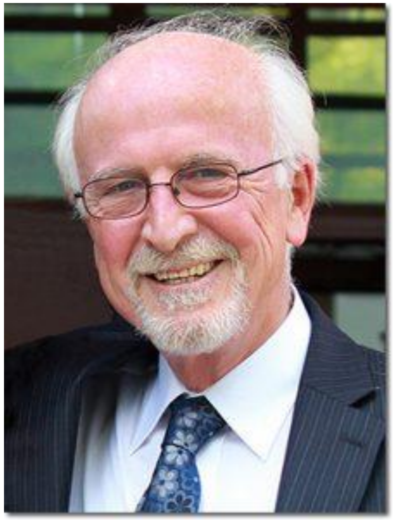
Computers & Chemical Engineering, Vol. 3, p. 573, 1979
Printed in Great Britain.
0098–1354/79/040573–01$02.00/0
Pergamon Press Ltd.

A UNIFIED APPROACH TO THE SYNTHESIS OF CONTROL STRUCTURES FOR COMPLEX CHEMICAL PLANTS

G. STEPHANOPOULOS* and Y. ARKUN
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, U.S.A.
and
M. MORARI
Department of Chemical Engineering, University of Wisconsin, Madison, WI 53706, U.S.A.

(*Received* 1 *December* 1979)

Research Inspiration #2



- **March 1995 – Imperial College London**
  - External PhD examiner for my first graduating PhD student – Dr. Katerina Papalexandri

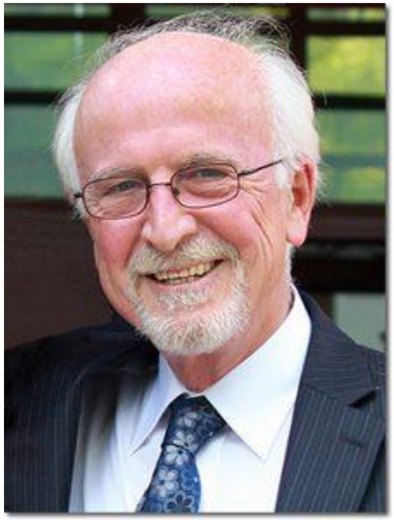
    "Flexibility and Controllability in the synthesis of mass/heat integrated process systems"

  - George offered encouragement/support and valuable advice

Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

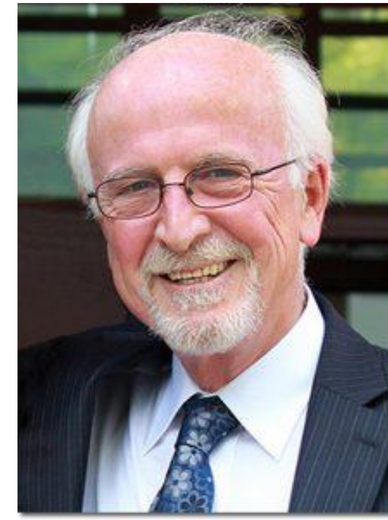
Research Inspiration #2



Encouragement to work on:

- Process Synthesis
- Process Intensification
- Interactions of design/control/operability
- Decision-making under uncertainty

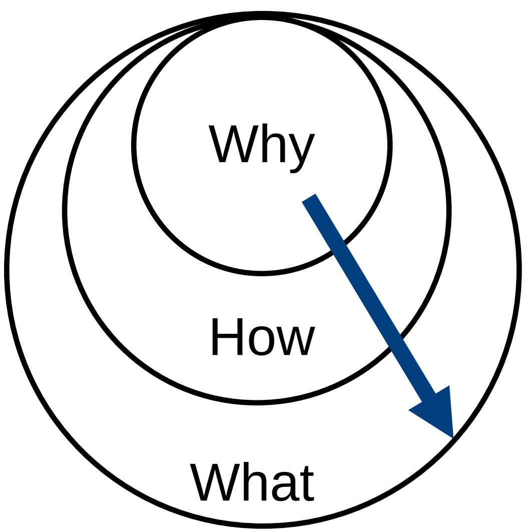
Research Inspiration #2



Valuable advice (for life):

- Research question is the key
- Have an area that you publish regularly (& you become well-known)
- Have a “risky” area with long-term impact

Importance of:

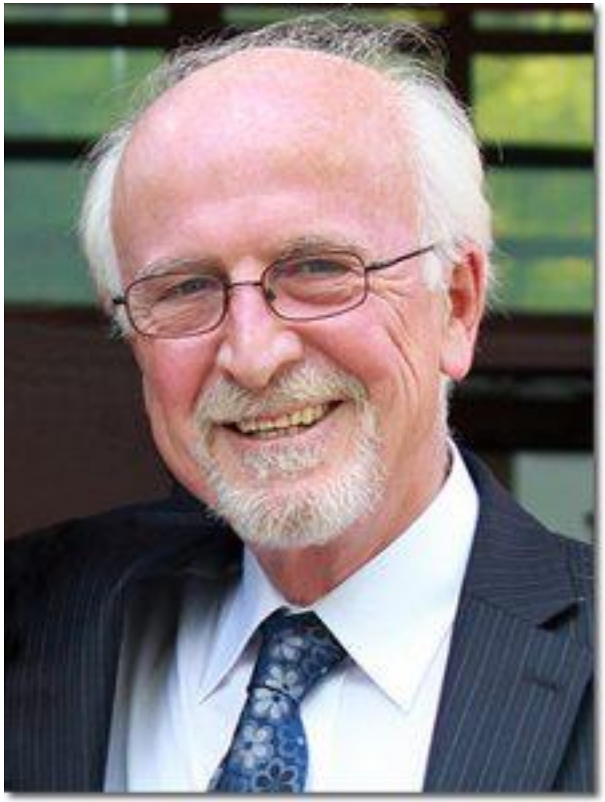


<u>From “why” to “how” to “what”:</u>

Strategy/Philosophy is critical to research innovation

(First - get the Big Picture right!)

Professor George Stephanopoulos



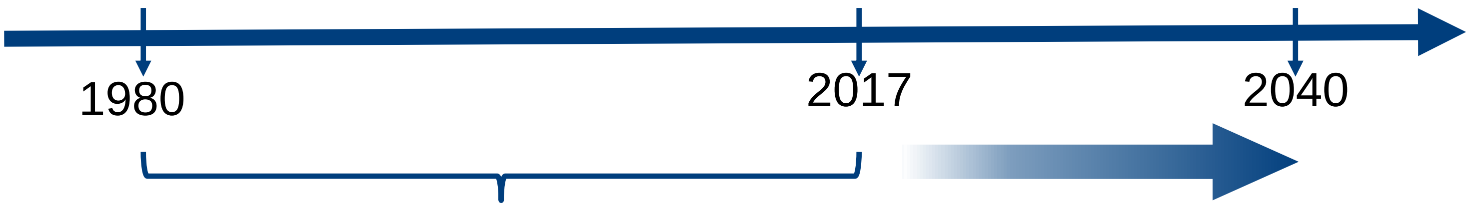
- Research inspiration & impact on my career
- Decision making under uncertainty in 2040

Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Decision-making under uncertainty in 2040

How far have we advanced between 1980-2017?



Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

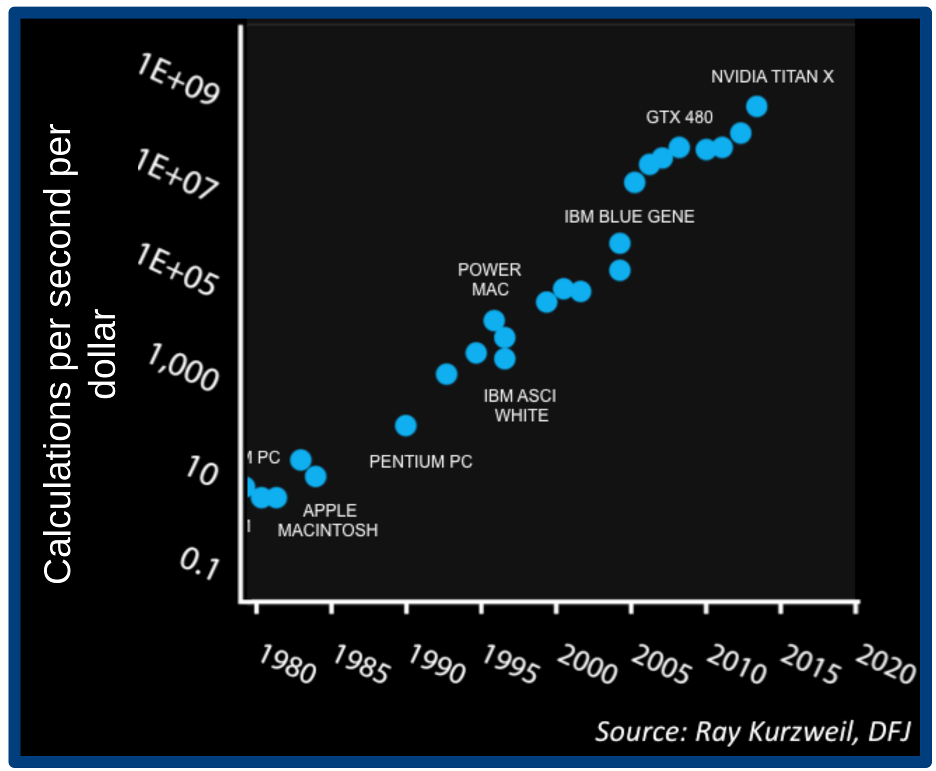
TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Decision-making under uncertainty in 2040

How far have we advanced between 1980-2017?

Advances in:

- Computing power
  - ×10E08 faster!



Decision-making under uncertainty in 2040

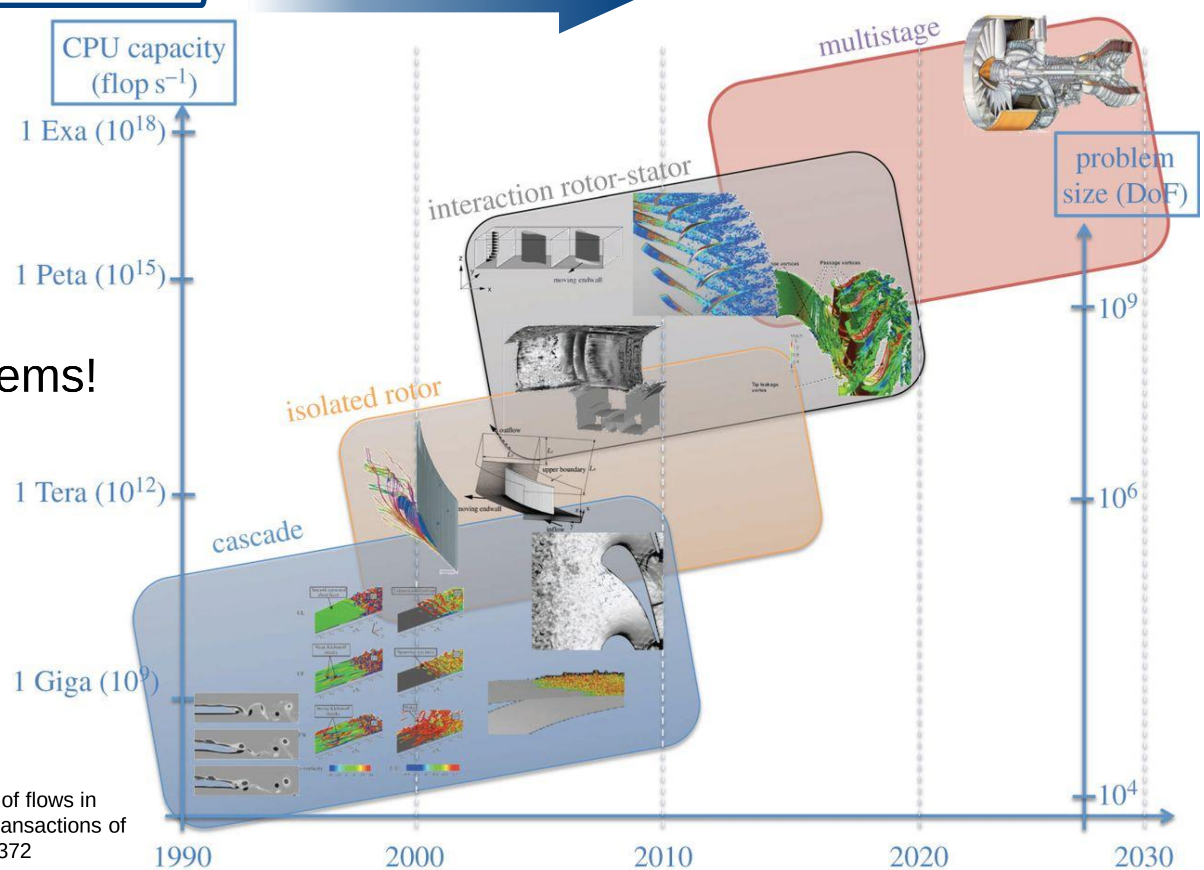
How far have we advanced between 1980-2017?

1980 2017 2040

Advances in:

- Computing power
  - ×10E08 faster!
- Simulation
  - ×10E04 bigger problems!

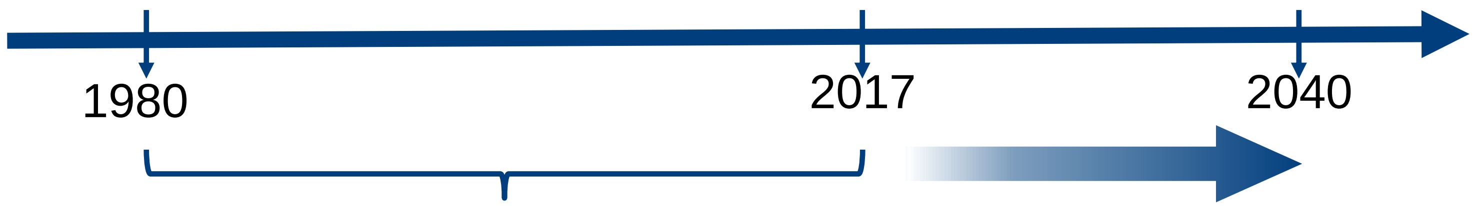
Simulation of flows in industrial compressors*



*Gourdain, N., Sicot, F., Duchaine, F., Gicquel, L. Large eddy simulation of flows in industrial compressors:A path from 2015 to 2035 (2014) Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372

Decision-making under uncertainty in 2040

How far have we advanced between 1980-2017?



Advances in:

- Computing power
  - ×10E08 faster!
- Simulation
  - ×10E04 bigger problems!

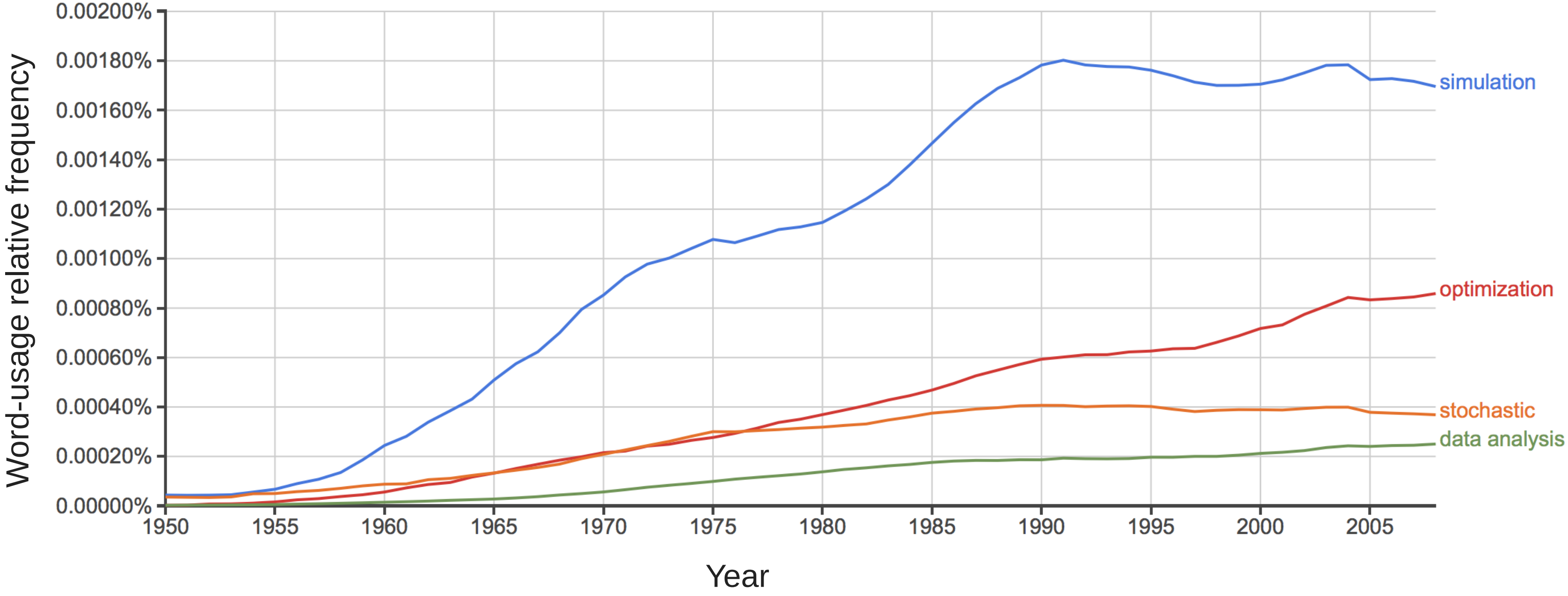
Similar trends in:

- Optimization
- Data Analytics

Decision-making under uncertainty in 2040

How far have we advanced between 1980-2017?

A word-usage relative frequency plot in Google Books



*Data is adopted from Google Ngrams

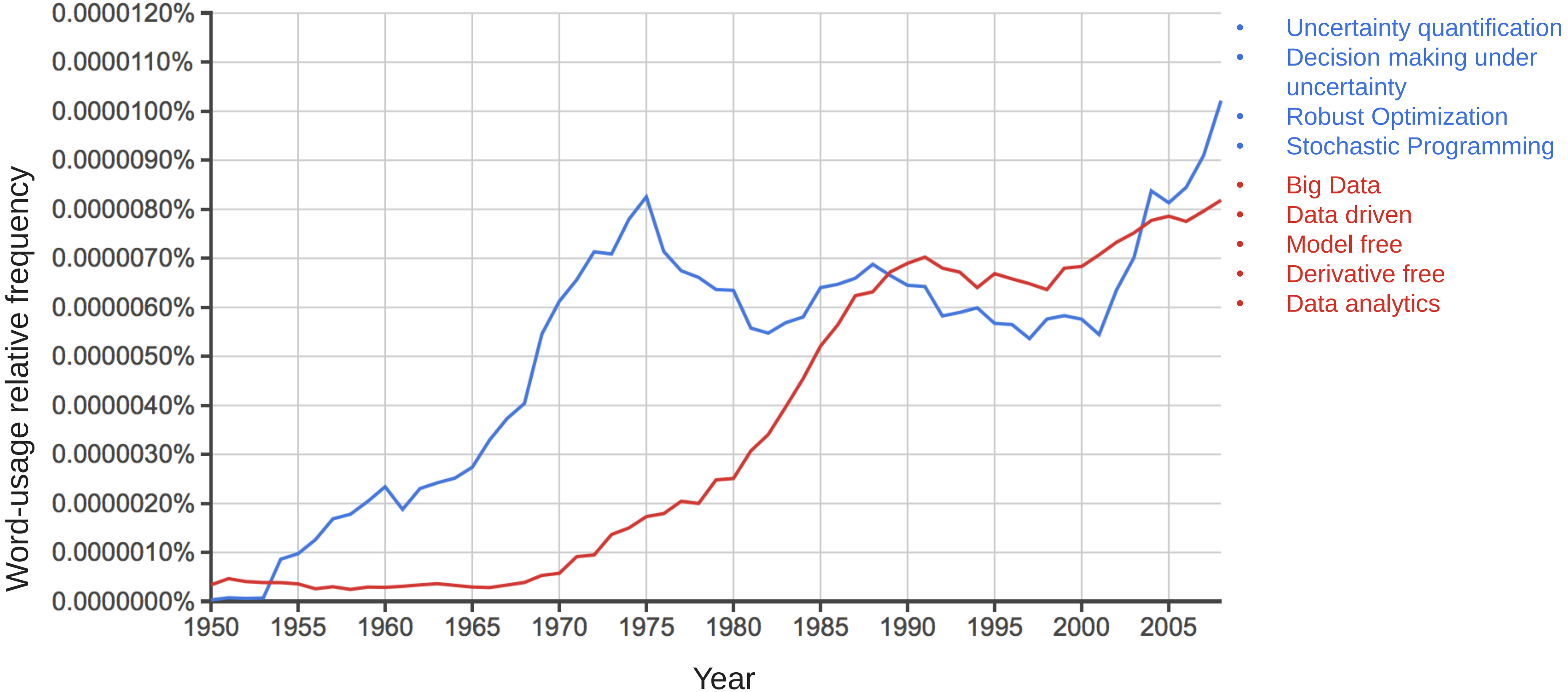
Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Decision-making under uncertainty in 2040

How far have we advanced between 1980-2017?

A word-usage relative frequency plot in Google Books



*Data is adopted from Google Ngrams

Decision-making under uncertainty in 2040

Adventure in Big Data Analytics or Opportunity for Model-based Optimization?

- Model -free ?
- Equation -free ?
- Variable -free ?
- Derivative -free ?

Myths & misconceptions?

Perception or Reality?

Decision-making under uncertainty in 2040

Adventure in Big Data Analytics or Opportunity for Model-based Optimization?

1. **Can we innovate / propose new designs with data only?**
   - Role of discrete 0-1 choices?
2. **Can Causality be fully expressed through data only?**
   - For example, biomedical applications
3. **Can we tackle decision making under uncertainty with data only?**
   - Two-stage stochastic/robust optimization

Motivating "thought" experiment

Two-stage decision-making under uncertainty

"here-and-now" **decisions:** Must be taken prior the realization of the **uncertainty**
"wait-and-see" **decisions:** Can be taken after the realization of the **uncertainty**

Application Areas:

1. Design and Scheduling:
   - **Design decisions:** "here-and-now"
   - **Scheduling decisions:** "wait-and-see"
   - **Demand:** uncertainty
2. Scheduling and Control:
   - **Scheduling decisions:** "here-and-now"
   - **Control decisions:** "wait-and-see"
   - **Process disturbances:** uncertainty
3. Design and Control
4. Facility location and transportation
5. Dynamic pricing and revenue management
6. Energy generation and distribution
7. Project management

Motivating "thought" experiment

Adjustable Robust Optimization

$$\min_{\mathbf{x}} \ \mathbf{c}^{\mathrm{T}}\mathbf{x} + \max_{\mathbf{u}\in U} \ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \ \mathbf{b}^{\mathrm{T}}\mathbf{y}$$

$$s.t. \quad \mathbf{Ax} \geq \mathbf{d}, \quad \mathbf{x} \in \mathbf{S}_{\mathbf{x}}$$

$$\Omega(\mathbf{x},\mathbf{u}) = \left\{ \mathbf{y} \in \mathbf{S}_{\mathbf{y}} : \mathbf{Wy} \geq \mathbf{h} - \mathbf{Tx} - \mathbf{Mu} \right\}$$

x: "here-and-now" decisions
y: "wait-and-see" decisions
u: uncertainty

Motivating "thought" experiment

Adjustable Robust Optimization

x: "here-and-now" decisions
y: "wait-and-see" decisions
u: uncertainty

Model-Based Approaches:

1. **One key approach – Assume linear decision rules:**

$$y = q + Qu$$

- Set y as an affine function of the uncertainty – simplifies the problem.
- Can be solved as *Static Robust Optimization*

2. **Generalized decision rules through multi-parametric programming:**

$$y = \begin{cases} Q_1 u + P_1 x + q_1 & if \quad G_1 u + H_1 x \le h_1 \\ Q_2 u + P_2 x + q_2 & if \quad G_2 u + H_2 x \le h_2 \\ & \vdots \\ Q_n u + P_n x + q_n & if \quad G_n u + H_n x \le h_n \end{cases}$$

- Solve the lower level problem **multi-parametrically** considering **u** and **x** as parameters
- Arrive to a set of (exact) affine decision rules valid for the whole feasible space of **u** and "here-and-now" decisions **x**.
- Use **M-POP** to get the **exact global** solution

Motivating “thought” experiment

Adjustable Robust Optimization

x: “here-and-now” decisions
y: “wait-and-see” decisions
u: uncertainty

Data-Driven Approaches:

3. Surrogate Modeling and Optimization within ARGONAUT

- Development of surrogate approximations that correlate input data to the problem objective.
- Use of **derivative-based** global optimization techniques to solve the optimization problem.

4. Machine Learning via Support Vector Machines

- Used to express **y** as an explicit function of **x** and **u**.
- **ϵ**-SVR: Supervised learning algorithm.
- Kernel functions transform **nonlinear data** to a **higher dimensional space** where there is a **linear** explanation of the input data.
- Kernel function used:

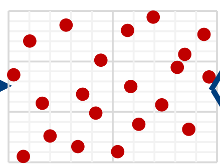
  $e^{-\gamma|x-x_i|^2}$, Parameters: γ (& C, ϵ)

Motivating "thought" experiment

Adjustable Robust Optimization

Sampling and Output Handling – Method 3

of inputs
Bounds



Latin Hypercube Design (LHD)

x, u

"Unsupervised"

- Solution via GAMS

$$\min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y}$$
$$s.t.\quad \mathbf{Ax}\geq\mathbf{d},\quad \mathbf{x}\in\mathbf{S}_x$$
$$\Omega(\mathbf{x},\mathbf{u})=\{\mathbf{y}\in\mathbf{S}_y:\mathbf{Wy}\geq\mathbf{h}-\mathbf{Tx}-\mathbf{Mu}\}$$

x

"Supervised"

- Solution via B-POP

$$\max_{\mathbf{u}\in U}\ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y}$$
$$s.t.\quad \mathbf{Ax}\geq\mathbf{d},\quad \mathbf{x}\in\mathbf{S}_x$$
$$\Omega(\mathbf{x},\mathbf{u})=\{\mathbf{y}\in\mathbf{S}_y:\mathbf{Wy}\geq\mathbf{h}-\mathbf{Tx}-\mathbf{Mu}\}$$

objective

objective

Surrogate Modeling and Optimization (ARGONAUT)

x optimal

x: "here-and-now" decisions
y: "wait-and-see" decisions
u: uncertainty

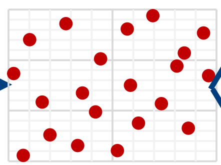
Motivating "thought" experiment

Adjustable Robust Optimization

Sampling and Output Handling – Method 4

of inputs

Bounds



Latin Hypercube Design (LHD)

x, u

"Unsupervised"

- Solution via GAMS

$$\min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y}$$
$$s.t.\quad \mathbf{Ax}\geq\mathbf{d},\quad \mathbf{x}\in\mathbf{S}_x$$
$$\Omega(\mathbf{x},\mathbf{u})=\left\{\mathbf{y}\in\mathbf{S}_y : \mathbf{Wy}\geq\mathbf{h}-\mathbf{Tx}-\mathbf{Mu}\right\}$$

y*

x

"Supervised"

- Solution via B-POP

$$\max_{\mathbf{u}\in U}\ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y}$$
$$s.t.\quad \mathbf{Ax}\geq\mathbf{d},\quad \mathbf{x}\in\mathbf{S}_x$$
$$\Omega(\mathbf{x},\mathbf{u})=\left\{\mathbf{y}\in\mathbf{S}_y : \mathbf{Wy}\geq\mathbf{h}-\mathbf{Tx}-\mathbf{Mu}\right\}$$

u*, y*

Machine Learning via Support Vector Machines

y as a function of x and u

x: "here-and-now" decisions
y: "wait-and-see" decisions
u: uncertainty

Motivating "thought" experiment

Instances:

	Instance 1	Instance 2	Instance 3
Linear/Non-Linear	Linear	Linear	Non-linear
# of "here-and-now" variables	2 - continuous	3 - continuous 3 - binary	1 - continuous 2 - binary
# of "wait-and-see" variables	2 - continuous	9 - continuous	3 - continuous
# of sampling points used by the data-driven techniques	41 – Unsupervised 21 - Supervised	728 – Unsupervised 488 - Supervised	244 – Unsupervised 124 - Supervised

Motivating "thought" experiment

Results:

Value of the objective function:

	Instance 1	Instance 2	Instance 3
Exact	451	30.536	7,320
Affine rules	451	33,680	----

Affine rules:

$$y = q + Qu$$

Exact – Multi-parametric approach:

$$y = \begin{cases} Q_1 u + P_1 x + q_1 & if \quad G_1 u + H_1 x \le h_1 \\ Q_2 u + P_2 x + q_2 & if \quad G_2 u + H_2 x \le h_2 \\ & \vdots \\ Q_n u + P_n x + q_n & if \quad G_n u + H_n x \le h_n \end{cases}$$

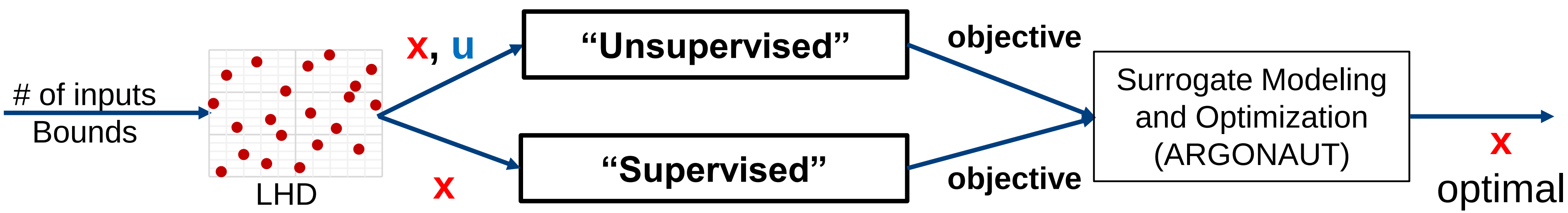
Motivating "thought" experiment

Results:

Value of the objective function:

	Instance 1	Instance 2	Instance 3
Exact	451	30.536	7,320
Affine rules	451	33,680	----
"Unsupervised" (ARGONAUT)	860	Infeasible	Infeasible
"Supervised" (ARGONAUT)	452.8	Infeasible	8,150

Surrogate Modeling and Optimization (ARGONAUT):



Motivating “thought” experiment

Results:

Value of the objective function:

	Instance 1	Instance 2	Instance 3
Exact	451	30.536	7,320
Affine rules	451	33,680	----
“Unsupervised” (ARGONAUT)	860	Infeasible	Infeasible
“Supervised” (ARGONAUT)	452.8	Infeasible	8,150

- Approximations (such as Affine rules) may be sub-optimal
- Affine rules cannot be derived for non-linear problems – Instance 3
- Data-driven techniques (could) arrive to near optimal solutions only for **continuous** problems with **“Supervised” data**
- **“Unsupervised” data** resulted in **infeasible solutions** for mixed-integer problems

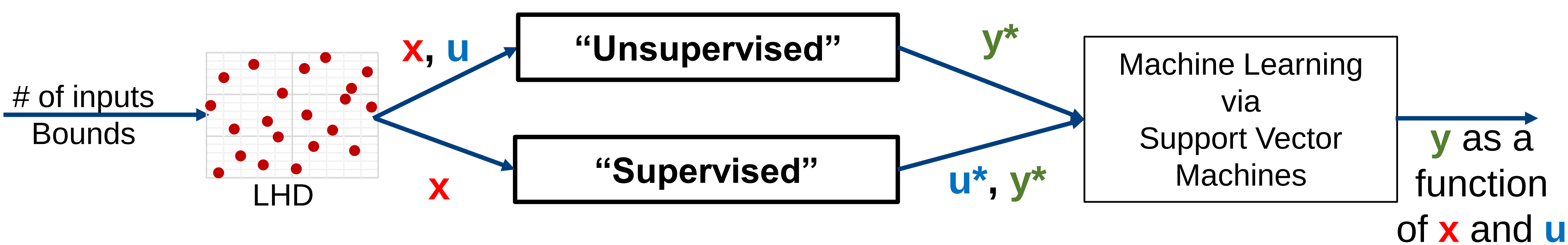
Motivating "thought" experiment

Results of the Machine Learning method:

		"Unsupervised"				"Supervised"			
		γ	C	ϵ	RMSE	γ	C	ϵ	RMSE
Instance 1	y_1	-	-	-	-	0.0078	1024	0.1	0.0848
	y_2	-	-	-	-	0.0625	8	0.1	0.0318
Instance 2	y_{11}	0.0313	1024	0.1	0.1013	0.1250	4	0.1	0.0051
	y_{12}	0.0313	1024	0.1	0.0981	0.0625	64	0.1	0.1029
	y_{13}	0.0313	512	0.1	0.1101	0.0625	512	0.1	0.0738
	y_{21}	-	-	-	-	0.0020	512	0.1	0.0916
	y_{22}	0.0313	512	0.1	0.1117	0.1250	16	0.1	0.0475
	y_{23}	0.0625	512	0.1	0.0996	0.0039	512	0.1	0.0259
	y_{31}	0.0313	512	0.1	0.1036	0.1250	128	0.1	0.1148
	y_{32}	0.0313	1024	0.1	0.1001	0.1250	32	0.1	0.0073
	y_{33}	0.0313	1024	0.1	0.0930	0.0625	32	0.1	0.0235
Instance 3	y_1	-	-	-	-	0.2500	32	0.1	0.0939
	y_2	-	-	-	-	8	128	0.1	0.5275
	y_3	-	-	-	-	0.1250	32	0.1	0.2857

Parameters of the function of **y** in terms of **x** and **u**

Root-Mean-Square Error



Motivating "thought" experiment

Results of the Machine Learning method:

		"Unsupervised"				"Supervised"			
		γ	C	ϵ	RMSE	γ	C	ϵ	RMSE
Instance 1	y_1	-	-	-	-	0.0078	1024	0.1	0.0848
	y_2	-	-	-	-	0.0625	8	0.1	0.0318
Instance 2	y_{11}	0.0313	1024	0.1	0.1013	0.1250	4	0.1	0.0051
	y_{12}	0.0313	1024	0.1	0.0981	0.0625	64	0.1	0.1029
	y_{13}	0.0313	512	0.1	0.1101	0.0625	512	0.1	0.0738
	y_{21}	-	-	-	-	0.0020	512	0.1	0.0916
	y_{22}	0.0313	512	0.1	0.1117	0.1250	16	0.1	0.0475
	y_{23}	0.0625	512	0.1	0.0996	0.0039	512	0.1	0.0259
	y_{31}	0.0313	512	0.1	0.1036	0.1250	128	0.1	0.1148
	y_{32}	0.0313	1024	0.1	0.1001	0.1250	32	0.1	0.0073
	y_{33}	0.0313	1024	0.1	0.0930	0.0625	32	0.1	0.0235
Instance 3	y_1	-	-	-	-	0.2500	32	0.1	0.0939
	y_2	-	-	-	-	8	128	0.1	0.5275
	y_3	-	-	-	-	0.1250	32	0.1	0.2857

- Using "Unsupervised" data - we cannot form expression for **y**
- Using "Supervised" data - expressions for **y** can be obtained for all cases (and are accurate)

Decision-making under uncertainty in 2040 – some remarks

- Data are important and useful provided that ..
  - The right type of data ..
  - The right amount of data …
  - The right timing of data …

- “Model-free”, “Variable-free”, “Equation-free” or “derivative-free” - a “myth”
  - A ‘model’ is always generated (with ‘variables’, ‘equations’ and use of ‘derivatives’!! – *within an algorithm*)

- Innovation/Design/Novel type of decisions **cannot** be obtained only by data!
  - Knowledge is essential

“Intelligent” data is the key

Decision-making under uncertainty in 2040 – some remarks

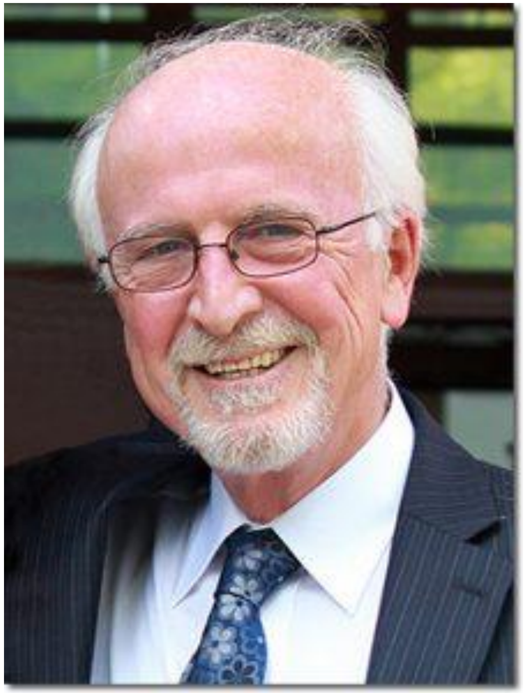
- Importance of
  - Algorithms
  - Hypothesis testing
- Big-Data Analytics
  - "virtual reality" modeling
  - "training" the algorithm (to create 'intelligence')
- Model-based Optimization
  - "reality" modelling
  - "training" the model (to create 'intelligence')
- Optimization is central in both!
  (Parameter estimation)
- **Hybrid approach** (?)

Data + **Models** + **Algorithms** → **Solution**

Artie McFerrin Department of CHEMICAL ENGINEERING TEXAS A&M UNIVERSITY

TEXAS A&M ENERGY INSTITUTE TEXAS A&M UNIVERSITY

Professor George Stephanopoulos



Happy 70th Birthday!
&
Happy Retirement!

Thank you for being such an Inspirational Intellectual Leader to all of us!!