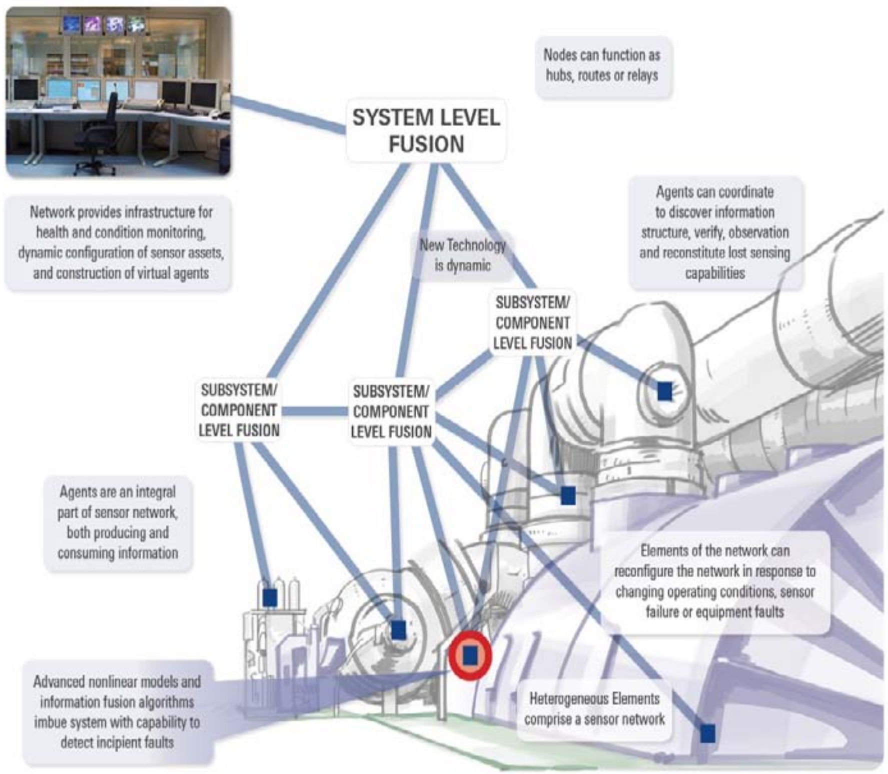



An Information-theoretic Framework and Self-organizing Sensor Network for Power Plant Equipment Monitoring

Crosscutting Research Kickoff Meeting

March 27, 2012



CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

Project Objective

Develop an information theoretic sensing and control framework that encompasses distributed software agents and computational methods to maximize the collection, transmission, aggregation, and conversion of data to actionable information for monitoring, diagnosis, prognosis and control of power plants.

A central goal is to enable (1) the extraction of all relevant information from sensor data, and (2) use the information obtained to enhance condition-based monitoring and improve operational control and decision-making at various levels of the control hierarchy.




Project Philosophy

To the extent that plant variables are not independent, information theory can be used to analyze and characterize the dependence beyond the simple notion of correlation.

The variability in temporal patterns of an observed quantity can be viewed as a signaling message, a flow of information about itself to all other parts of the system that are "listening."

View systems as networks of information transfer!

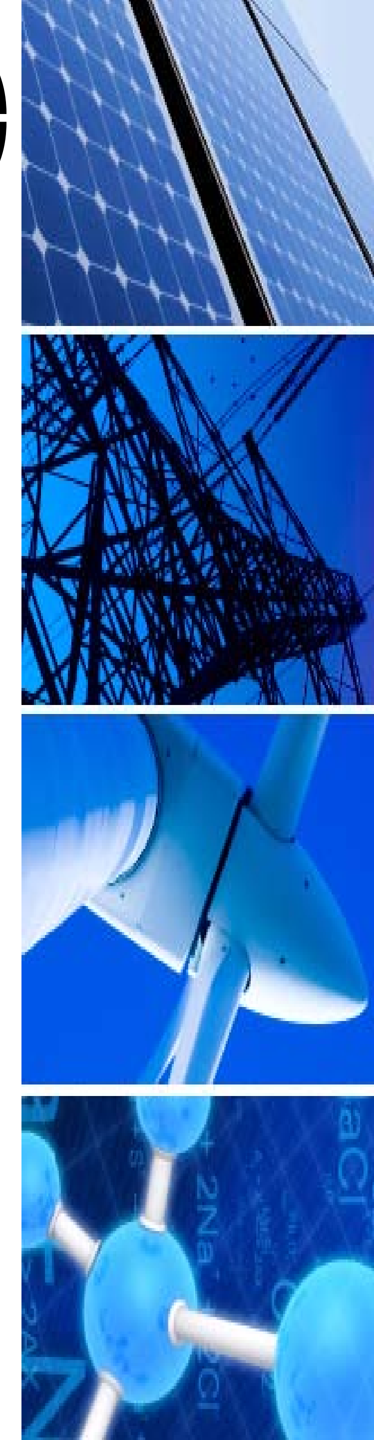
Technical Objectives

1. ***Develop an intelligent agent-based information-theoretic architecture*** for advanced power plant applications.
2. ***Develop computational algorithms*** to maximize the collection, transmission, aggregation, and conversion of data into actionable information for monitoring, diagnosis, prognosis and control of the power plant.
3. ***Evaluate*** the effectiveness of these ***computational algorithms*** for maximizing information content from power plant data through an ***integrated hardware-in-the-loop simulation test bed***.

Innovations

1. ***Active probing capability:*** Enriches the information content of the available observations by addressing the intrinsic trade-off between estimation and control within an information-theoretic context
2. ***Virtual sensors:*** Provide a capability to discover the correlative structure of the available observations and to fuse information from disparate sources
3. ***Compressive sensing algorithms:*** Feature extraction and minimal representations in a network setting
4. ***Distributed intelligent agents:*** Decentralized processing to maximize available computational resources.

Testing and validation will be performed in a hardware-in-the loop simulation environment



Project Scope

- Focus on monitoring, diagnosis and prognosis of power plant equipment and processes:
  - *Will not directly address control but will accommodate its future implementation.*
  - *Work will focus on fossil fuel power plants.*
- Includes monitoring and diagnosis of the information network and instrumentation.

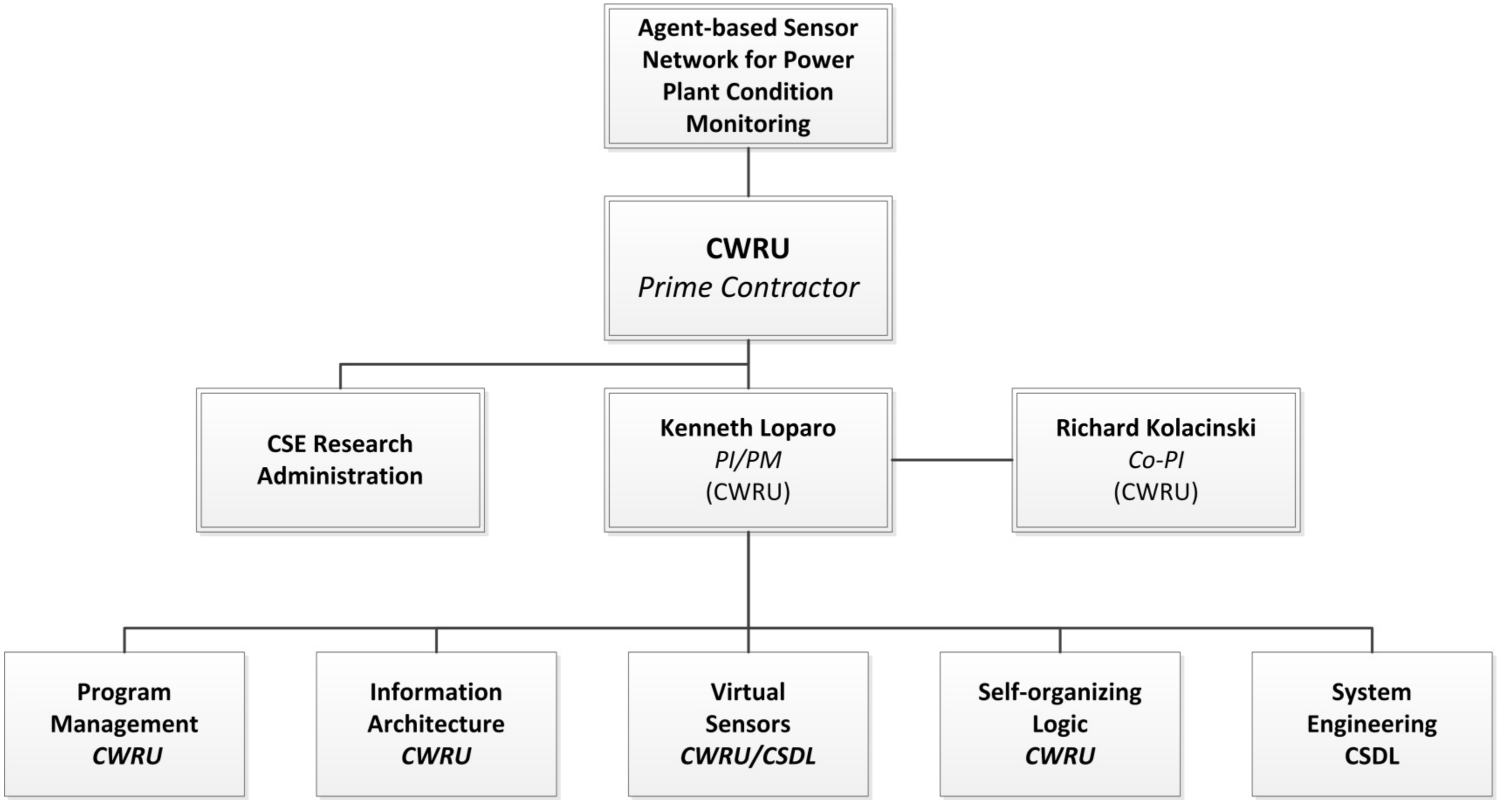
Case Western Reserve University | Great Lakes Energy Institute

Draper Laboratory

Tasks

- Work will be organized into three phases:
  - ***Phase 1*** *will focus on requirements elicitation and analysis.*
  - ***Phase 2*** *will focus on algorithm development.*
  - ***Phase 3*** *will focus on integration of a demonstration system.*
- Five tasks:
  - ***Task 1:*** *Program Management*
  - ***Task 2:*** *Information Architecture*
  - ***Task 3:*** *Virtual Sensing*
  - ***Task 4:*** *Self-organizing Logic*
  - ***Task 5:*** *System Engineering*

Organizational Work Breakdown Structure



- Task leadership shared between Case Western Reserve University and C. S. Draper Laboratory.
- Leverage each institution’s particular strengths.

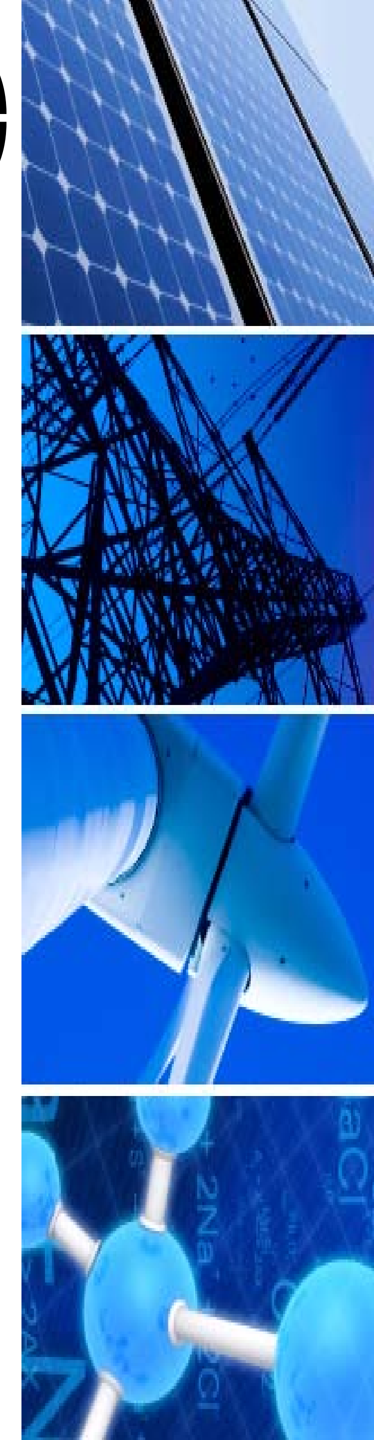
CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

Work Breakdown Structure

Self-organizing Sensor Network using Information Theoretic Principles

- **1. Program Management**
  - 1.1 Project Start-up
  - 1.2 Meetings and Review
  - 1.3 Project Reporting
  - 1.4 Project Control
  - 1.5 Project Administration
  - 1.6 Project Support
  - 1.7 Project Close-out
- **2. Information Architecture**
  - 2.1 Analysis
  - 2.2 Development
  - 2.3 Implementation
- **3. Virtual Sensors**
  - 3.1 Analysis
  - 3.2 Development
  - 3.3 Implementation
- **4. Self-Organization**
  - 4.1 Analysis
  - 4.2 Development
  - 4.3 Implementation
- **5. System Engineering**
  - 5.1 Analysis
  - 5.2 Development
  - 5.3 Integration and Testing

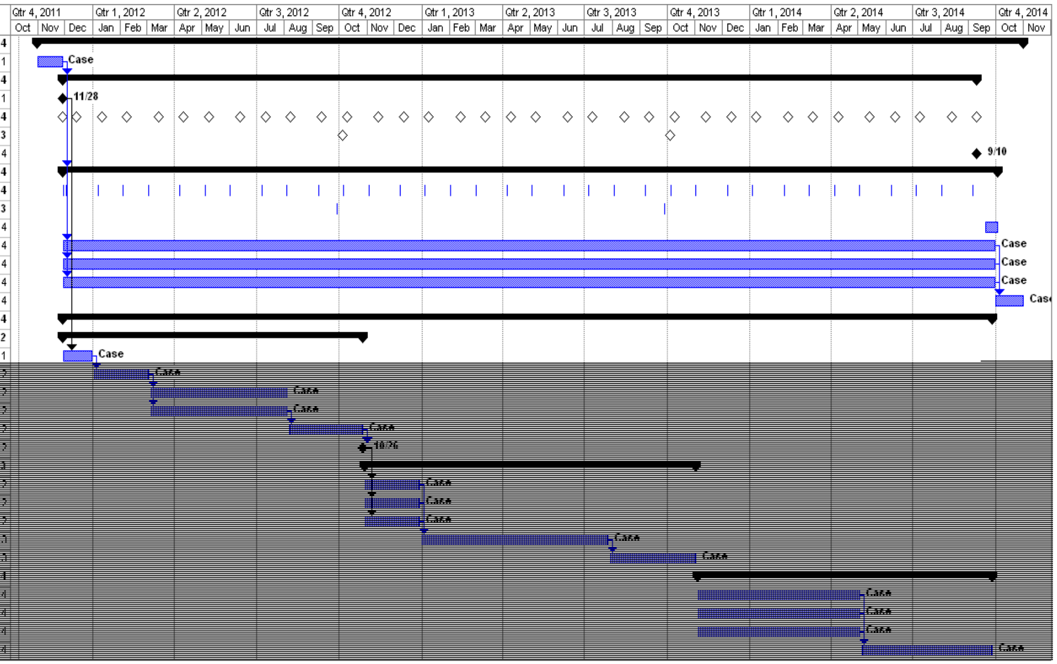


Case Western Reserve University | Great Lakes Energy Institute

Draper Laboratory

Information Architecture

ID		Task Name	Duration	Start	Finish
1		**Project Management and Planning**	**785 days**	**Mon 10/31/11**	**Fri 10/31/14**
2		Project Start-up	21 days	Mon 10/31/11	Mon 11/28/11
3		**Meetings and Review**	**726 days**	**Mon 11/28/11**	**Wed 9/10/14**
4		Kick-off Meeting	0 days	Mon 11/28/11	Mon 11/28/11
5		**Monthly Review Meeting**	**726 days**	**Mon 11/28/11**	**Wed 9/10/14**
41		**Annual Review Meetings**	**260 days**	**Fri 10/5/12**	**Fri 10/4/13**
44		Close-out Meeting	0 days	Wed 9/10/14	Wed 9/10/14
45		**Project Reporting**	**743 days**	**Tue 11/29/11**	**Thu 10/2/14**
46		**Monthly Reporting**	**724 days**	**Tue 11/29/11**	**Fri 9/5/14**
82		**Annual Report**	**261 days**	**Fri 9/28/12**	**Fri 9/27/13**
85		Final Report	10 days	Fri 9/19/14	Thu 10/2/14
86		Project Control	740 days	Tue 11/29/11	Mon 9/29/14
87		Project Administration	740 days	Tue 11/29/11	Mon 9/29/14
88		Project Support	740 days	Tue 11/29/11	Mon 9/29/14
89		Project Close-out	24 days	Tue 9/30/14	Fri 10/31/14
90		**Information Architecture**	**739 days**	**Tue 11/29/11**	**Fri 9/26/14**
91		**Analysis**	**239 days**	**Tue 11/29/11**	**Fri 10/26/12**
92		Plant Equipment Survey	24 days	Tue 11/29/11	Fri 12/30/11
93		Plant Sensors and Phenomena Su	45 days	Mon 1/2/12	Fri 3/2/12
94		Information Structure of Observati	110 days	Mon 3/5/12	Fri 8/3/12
95		Sensor/Transmission Characteriz	110 days	Mon 3/5/12	Fri 8/3/12
96		Metrics and Objectives Specificati	60 days	Mon 8/6/12	Fri 10/26/12
97		Network Performance Objectives	0 days	Fri 10/26/12	Fri 10/26/12
98		**Development**	**265 days**	**Mon 10/29/12**	**Fri 11/1/13**
99		Output Requirement Specification	45 days	Mon 10/29/12	Fri 12/28/12
100		Observation Process Requirement	45 days	Mon 10/29/12	Fri 12/28/12
101		Observation Objective Function De	45 days	Mon 10/29/12	Fri 12/28/12
102		Information Constraint and Partion	150 days	Mon 12/31/12	Fri 7/26/13
103		Automated Comm. Architecture Ge	70 days	Mon 7/29/13	Fri 11/1/13
104		**Implementation**	**235 days**	**Mon 11/4/13**	**Fri 9/26/14**
105		Software Development	130 days	Mon 11/4/13	Fri 5/2/14
106		Unit/Functional Testing	130 days	Mon 11/4/13	Fri 5/2/14
107		Use case/Scenario Generation	130 days	Mon 11/4/13	Fri 5/2/14
108		Architecture Testing and Evaluatio	105 days	Mon 5/5/14	Fri 9/26/14



- Develop and implement an information theoretic framework for heterogeneous sensor networks:
  - *Optimize extraction and flow of information from sensors and related objects to relevant sensor processing, aggregation, and fusion processes*
  - *Subject to system requirements for bandwidth, robustness, and fault tolerance.*

Information

- **Information is the amount of "surprise" contained in data;**
  - Data is not information!
  - Information is related to new knowledge – data that tells you what you already know is not informative.
- **Example – English Language:**
  - A "u" following a "q" is not surprising, and we say that the "u" contains *no information*,
  - While not definite, an "h" following a "t" is not a complete surprise and we say that the "h" does not contain *full information*,
  - A "g" following a "q" would be very surprising and thus we say that the "g" contains *full information*.

CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

Information Measures

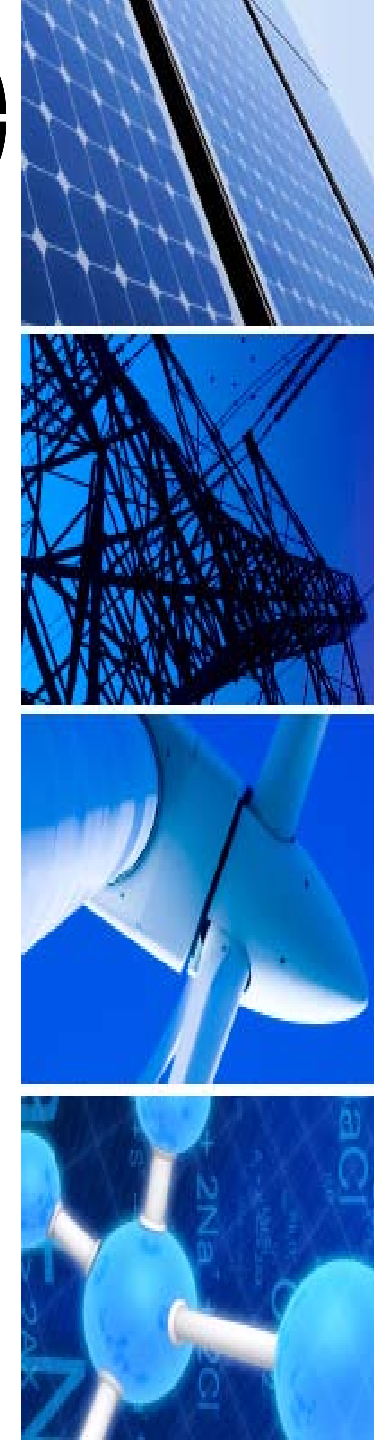
- The basic measure of information is entropy:

$$H(X) = -\sum_{x \in X} p(x)\log_2(p(x))$$

  - *Quantifies the unpredictability of a random variable-average number of bits required to describe a random variable X.*

- Extensions characterize relationships between multiple variables and distributions:
  - Joint Entropy: $H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2 p(x,y)$
  - Conditional Entropy: $H(X|Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2 p(x|y)$
  - Relative Entropy: $D(p\|q) = \sum_{x\in X} p(x)\log_2 \frac{p(x)}{q(x)}$
  - Mutual Information:

$$\begin{aligned} I(X;Y) &= \sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2 \frac{p(x,y)}{p(x)p(y)} \\ &= D\left(p(x,y)\|p(x)p(y)\right) \end{aligned}$$



Chain Rule for Entropy Calculations

- ***Chain Rule for Entropy (Bivariate):***

$$H(X,Y) = H(X) + H(Y|X)$$

  - Note, in general, $H(X|Y) \neq H(Y|X)$ though
  - .

$$H(X) - H(X|Y) = H(Y) - H(Y|X)$$

  - An immediate corollary is

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z)$$

- ***Chain Rule for Entropy (Multivariate):***

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i-1}, \ldots, X_1)$$

Chain Rule for Entropy Calculations (cont.)

- Conditional Relative Entropy:

$$D\left(p(y|x)\|q(y|x)\right) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \frac{p(y|x)}{q(y|x)}$$

- Chain Rule for Relative Entropy:

$$D\left(p(x,y)\|q(x,y)\right) = D\left(p(x)\|q(x)\right) + D\left(p(y|x)\|q(y|x)\right)$$

- Conditional Mutual Information:

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$$

- Chain Rule for Mutual information:

$$I(X_1, X_2, \ldots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y|X_{i-1}, \ldots, I_1)$$

CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

Estimation

- Can formulate estimators with alternative goals:
  - ***Minimum Mean Squared Error*** *estimator,* $\bar{f}_g$ *:*

$$\sigma^{g,\bar{f}_g}(\tilde{x}_k) = \inf_{f \in \mathrm{F}} E\left\{ \left\| \tilde{x}_k^{g,f} \right\|^2 \right\}$$

  - ***Minimum Error/Observation Information*** *estimator,* $\tilde{f}_g$*:*

$$I\left(x_k^{g,\tilde{f}_g} ; Y_k^g \right) = \inf_{f \in \mathrm{F}} I\left(x_k^{g,f} ; Y_k^g \right)$$

  - ***Minimum Error Entropy*** *estimator,* $\hat{f}_g$*:*

$$h\left(\tilde{x}_k^{g,\hat{f}_g} \right) = \inf_{f \in \mathrm{F}} h\left(x_k^{g,f} \right)$$

- For Linear Gaussian systems estimators are equivalent;
  - *Separation principle for LQG problem*
  - *Optimal filter is the Kalman Filter*
- In general, there is a probing effect and the appropriate control and filter pair (*g* , *f*) must be designed.

Case Western Reserve University | Great Lakes Energy Institute

Draper Laboratory

Control/Filter Design

- Estimation and Control objectives must be addressed simultaneously;
  - *Minimum Mean Square pair, (g^*, f^*) :*

$$E\left\{\left\|\tilde{x}_k^{g^*,f^*}\right\|^2\right\} = \inf_{g\in G,\, f\in F} E\left\{\left\|\tilde{x}_k^{g,f}\right\|^2\right\}$$

  - *Minimum Error Observation pair, $(g^{\#}, f^{\#})$;*

$$I\left(\tilde{x}_k^{g^{\#},f^{\#}}; Y_k^{g^{\#}}\right) = \inf_{g\in G,\, f\in F} I\left(\tilde{x}_k^{g,f}; Y_k^{g}\right)$$

  - *Minimum Error Entropy pair, $(g^{\%}, f^{\%})$;*

$$h\left(\tilde{x}_k^{g^{\%},f^{\%}}\right) = \inf_{g\in G,\, f\in F} h\left(\tilde{x}_k^{g,f}\right)$$

- Maximum Channel Transmittance control, *$g^{\$}$;*

$$I\left(x_k^{g^{\$}}; Y_k^{g^{\$}}\right) = \sup_{g\in G} I\left(x_k^{g}; Y_k^{g}\right)$$

  - *Provides an upper bound on information that can be extracted from observations by a filter,*
  - *If a sensor is a channel, this is the channel capacity.*

Entropy Performance Index

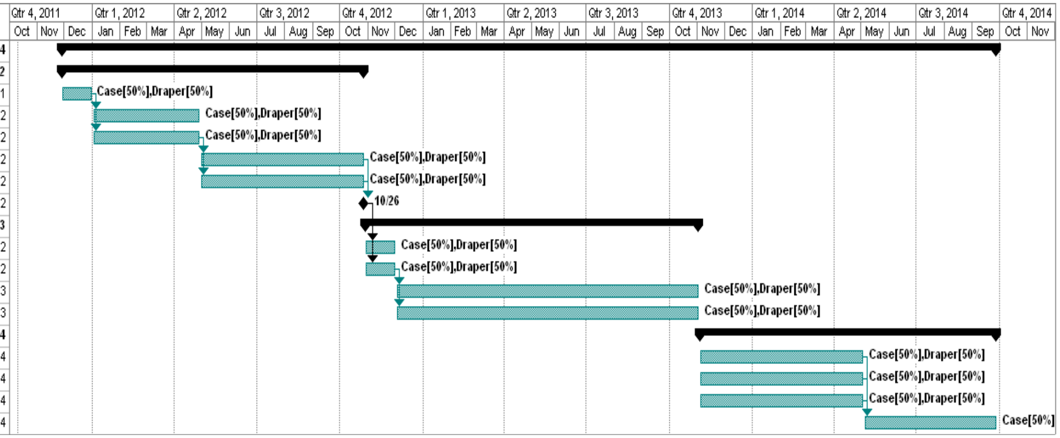
- The synthesis of a probing control g requires maximizing the sensor channel transmittance while minimizing the error/observation mutual information;
  - *These goals conflict!*
  - *Multiobjective optimization problem.*
- One approach is to maximize the difference, i.e.,

$$\sup_{g \in \mathrm{G}, f \in \mathrm{F}} W^{g,f} \triangleq \max_{g \in \mathrm{G}, f \in \mathrm{F}} \left[I^{g}\left(x_k ; Y_k\right) - I^{g,f}\left(\tilde{x}_k ; Y_k\right) \right]$$
$$= \sup_{g \in \mathrm{G}, f \in \mathrm{F}} \left[h^{g}\left(x_k\right) - h^{g,f}\left(\tilde{x}_k\right) \right]$$

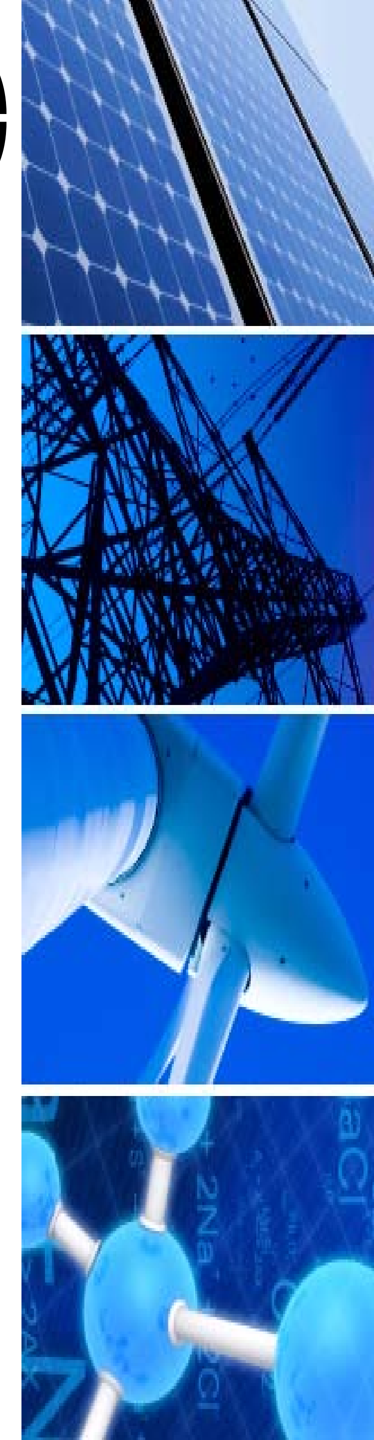
  - *$W^{g,f}$ is referred to as the **Entropy Performance Index**,*
  - *Measures the difference between the differential entropy of the input to the channel x_k^g and the differential entropy of the output from the channel $\tilde{x}_k^{g,f}$,*
  - *$W^{g,f} \geq 0$* for all “nontrivial” filters.

Virtual Sensors

ID		Task Name	Duration	Start	Finish
109		**Virtual Sensor/Sensor Processing**	**739 days**	**Tue 11/29/11**	**Fri 9/26/14**
110		**Analysis**	**239 days**	**Tue 11/29/11**	**Fri 10/26/12**
111		Monitoring Needs Elicitation	24 days	Tue 11/29/11	Fri 12/30/11
112		Virtual Sensor Requirements Flow	85 days	Mon 1/2/12	Fri 4/27/12
113		Information Processing Survey	85 days	Mon 1/2/12	Fri 4/27/12
114		Information Processing Methods S	130 days	Mon 4/30/12	Fri 10/26/12
115		ID Processing Capability Gaps	130 days	Mon 4/30/12	Fri 10/26/12
116		Virtual Sensor Methods	0 days	Fri 10/26/12	Fri 10/26/12
117		**Development**	**265 days**	**Mon 10/29/12**	**Fri 11/1/13**
118		Sensor Output Specification	25 days	Mon 10/29/12	Fri 11/30/12
119		Sensor Interface Specification	25 days	Mon 10/29/12	Fri 11/30/12
120		Information Processing Algorithm [	240 days	Mon 12/3/12	Fri 11/1/13
121		Virtual Sensor Logic and Interface	240 days	Mon 12/3/12	Fri 11/1/13
122		**Implementation**	**235 days**	**Mon 11/4/13**	**Fri 9/26/14**
123		Software Development	130 days	Mon 11/4/13	Fri 5/2/14
124		Unit/Functional Testing	130 days	Mon 11/4/13	Fri 5/2/14
125		Use-case/Scenario Generation	130 days	Mon 11/4/13	Fri 5/2/14
126		Virtual Sensor Test and Evaluatior	105 days	Mon 5/5/14	Fri 9/26/14



- Develop a functional link between the information-theoretic framework developed under Task 2 and the Self Organizing Capability developed under Task 4:
  - *Develop information extraction and processing algorithms intended for implementation within software agents.*
  - *Develop the basic tools for realizing the complementary principle of minimizing complexity by providing the wherewithal to incorporate feature extraction techniques for minimizing description lengths.*

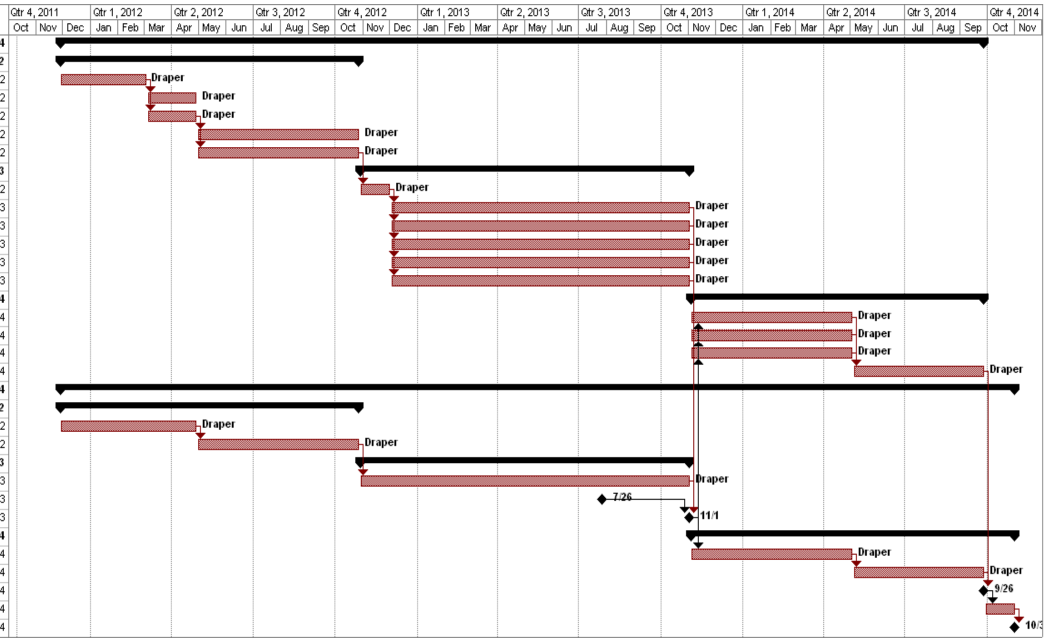


CASE WESTERN RESERVE UNIVERSITY | GREAT LAKES ENERGY INSTITUTE

DRAPER LABORATORY

Self-organizing Logic

ID		Task Name	Duration	Start	Finish
127		**Self-organizing Logic**	**739 days**	**Tue 11/29/11**	**Fri 9/26/14**
128		**Analysis**	**239 days**	**Tue 11/29/11**	**Fri 10/26/12**
129		Survey of Self-Organizing Approa	69 days	Tue 11/29/11	Fri 3/2/12
130		SW Agent Requirements	40 days	Mon 3/5/12	Fri 4/27/12
131		Select Candidate Self-org Approa	40 days	Mon 3/5/12	Fri 4/27/12
132		SW Agent Sim Environment	130 days	Mon 4/30/12	Fri 10/26/12
133		Metric Generation	130 days	Mon 4/30/12	Fri 10/26/12
134		**Development**	**265 days**	**Mon 10/29/12**	**Fri 11/1/13**
135		ID SW Agent Types	25 days	Mon 10/29/12	Fri 11/30/12
136		Agent Logic Development	240 days	Mon 12/3/12	Fri 11/1/13
137		Agent Comm. Protocols	240 days	Mon 12/3/12	Fri 11/1/13
138		Agent Data Interface Specificatior	240 days	Mon 12/3/12	Fri 11/1/13
139		System Test Procedure Developm	240 days	Mon 12/3/12	Fri 11/1/13
140		System Test Framework Develom	240 days	Mon 12/3/12	Fri 11/1/13
141		**Implementation**	**235 days**	**Mon 11/4/13**	**Fri 9/26/14**
142		Software Development	130 days	Mon 11/4/13	Fri 5/2/14
143		Unit/Functional Testing	130 days	Mon 11/4/13	Fri 5/2/14
144		Use-case/Scenario Generation	130 days	Mon 11/4/13	Fri 5/2/14
145		Agent Test and Evaluation	105 days	Mon 5/5/14	Fri 9/26/14
146		**System Engineering & Integration**	**764 days**	**Tue 11/29/11**	**Fri 10/31/14**
147		**Analysis**	**239 days**	**Tue 11/29/11**	**Fri 10/26/12**
148		Requirements Analysis	109 days	Tue 11/29/11	Fri 4/27/12
149		System Analysis	130 days	Mon 4/30/12	Fri 10/26/12
150		**Development**	**265 days**	**Mon 10/29/12**	**Fri 11/1/13**
151		System Design	265 days	Mon 10/29/12	Fri 11/1/13
152		PDR	0 days	Fri 7/26/13	Fri 7/26/13
153		CDR	0 days	Fri 11/1/13	Fri 11/1/13
154		**Implementation**	**260 days**	**Mon 11/4/13**	**Fri 10/31/14**
155		Integration and Assembly	130 days	Mon 11/4/13	Fri 5/2/14
156		System Testing	105 days	Mon 5/5/14	Fri 9/26/14
157		Demonstration System Complete	0 days	Fri 9/26/14	Fri 9/26/14
158		System Demonstration	25 days	Mon 9/29/14	Fri 10/31/14
159		Working Prototype	0 days	Fri 10/31/14	Fri 10/31/14



- Develop and implement a self-organizing, multi-agent communications and control architecture:
  - *Identification of disparate agents, their behaviors and interactions, and direct and stigmergic communications.*
  - *Elucidate organizing principles of system components necessary to achieve overall system objectives including maximizing plant reliability and enhancing information discovery.*

Case Western Reserve University | Great Lakes Energy Institute

Draper Laboratory

Milestones

ID	WBS Element	Description	Planned Complete Date	Actual Complete Date	Comments
1	1	Kick-off Meeting	28-Nov. 2011	27-Mar. 2012	
2	1	Close-out Meeting	31-Oct. 2014		
		Information Architecture Milestones			
3	2	Network Performance Objectives Specified	26-Oct. 2012		
		Virtual Sensor/Sensor Processing Milestones			
4	3	Virtual Sensor Methods Identified	26-Oct. 2012		
		System Integration Milestones			
5	5	Preliminary Design Review (PDR)	26-Jul. 2013		
6		Critical Design Review (CDR)	31-Oct. 2013		
7		Demonstration System Completed	26-Sept. 2014		
8		System Demonstrated	31-Oct. 2014