Overview

Property based Techniques for Process and Product Design

Dr. Mario Richard Eden
Mary & John H. Sanders Associate Professor
Department of Chemical Engineering
Auburn University

Department of Chemical and Biomolecular Engineering
Georgia Institute of Technology
February 3, 2010
My Background

• **Background**

• **Professional Experience**
  – Associate Professor, Auburn University (2008 – present)
  – Assistant Professor, Auburn University (2004 – 2008)
Where is Denmark?
A Few Facts about Denmark

Constitutional Monarchy

A little smaller than the state of Alabama (not including Greenland)

Population approximately 5500000.

National sport – SOCCER!

Where I moved to go to college

My hometown
• **Computer Aided Process Engineering**
  – Property prediction & CAMD for solvent selection/design
  – Process modeling and simulation

• **Process/ Product Synthesis and Design**
  – Develop novel efficient methods for emerging problems
  – Develop strategies for simultaneous solution
  – Systematic identification/generation of alternatives

• **Process Integration and Optimization**
  – Application of holistic methods to ensure sustainability
  – Fuels reforming and biorefinery optimization
Challenges and Motivation

Design Objectives - Identify:

- Important fruits (products)
- Optimal path to reach them
- Feasibility of the process
- Control structure
- Resource conservation strategies
- Environment, health and safety issues

Highly complex problems due to interactions between them!!
**General Design Problem 1:2**

**Process/ Product Synthesis/ Design Problem**

\[ F_{OBJ} = \min \{ C^T \cdot y + f(x) \} \]

s.t. \[ h_1(x, y) = 0 \]
\[ h_2(x, y) = 0 \]
\[ g_1(x) > 0 \]
\[ g_2(x, y) > 0 \]
\[ B \cdot y + C \cdot x > d \]

- **Objective Function**
- **Process/Product Model**
- **Equality and Inequality Constraints**
- **Structural Constraints**

**Conventional Solutions**

**A. Heuristic/ knowledge-based**
Satisfy only the constraints. Generates feasible solutions.

**B. Mathematical Optimization**
Solve the objective function including the process model.
General Design Problem 2:2

Process/ Product Synthesis/ Design Problem

\[ F_{\text{OBJ}} = \min \{ C^T \cdot y + f(x) \} \]
\[ \text{s.t.} \quad h_1(x, y) = 0 \]
\[ h_2(x, y) = 0 \]
\[ g_1(x) > 0 \]
\[ g_2(x, y) > 0 \]
\[ B \cdot y + C \cdot x > d \]

Hybrid Solution Approach

I. Define search space through heuristic or knowledge-based approach (A).

II. Solve well defined optimization problem (B).

IMPORTANT

Regardless of problem type a process model is needed and it is the model type and validity ranges that defines the application range of the solution.
Constitutive Model

Choice of a constitutive model implicitly defines the search space

Do we know enough to derive a single model?

Can we solve simulation and/or optimization problems with multiple constitutive models representing the same variable?
1. Service Role for Process Model
2. Service & Advice Role for Process Model
Roles of Property Models

Compounds needed only by property model
Service, Advice & Solve Role for Process Model

New approach: Problem size reduced
Targeted Design Techniques

- **Reverse Problem Formulation Methodology**
  - Computationally efficient solution strategy
  - Targets optimal solution prior to complex calculations
  - Capable of identifying all alternatives
  - Framework enables optimum use of all model details
  - Relieves inherently iterative nature of design
Reverse Problem Formulation

**Process Model**
- Balance and Constraint Equations (Mass, Energy, Momentum)
- Constitutive Equations (Phenomena model - Function of Intensive Variables)

**Balance and Constraint Equations**
Identification of design targets by solution of decoupled model

**DESIGN TARGETS (Constitutive Variables)**
Match design targets and identify corresponding intensive variables

**As long as the targets are matched, the balance equations are also satisfied and do NOT need to be solved again**
Why Design Based on Properties?

- Many processes driven by properties NOT components
- Performance objectives often described by properties
- Often objectives can not be described by composition
- Product/molecular design is based on properties
- Insights hidden by not integrating properties directly

Property Clusters

- Extension to existing composition based methods
- Reduces dimensionality of problem
- Enables visualization of problem
- Property estimation in molecular design via GC
- Unifying framework for simultaneous solution
• **Property Clustering**
  - Conserved surrogate properties described by property operators, which have linear mixing rules, even if the operators themselves are nonlinear.
  - Deconstructs the design problem into a Euclidean vector in the cluster domain and a scalar called the Augmented Property Index (AUP).

**Clustering Algorithm**

- **Linearization of Property Expression**
  \[ \psi_j(p) = \rho_j \]
- **Property Operator**
  \[ \psi_{jm} = \sum_{i=1}^{n} x_i \cdot \psi_j \]
- **Normalization**
  \[ \Omega_j = \frac{\psi_j}{\psi_j^{ref}} \]
- **Augmented Property Index**
  \[ AUP = \sum_{j=1}^{m} \Omega_j \]
- **Property Cluster**
  \[ C_j = \frac{\Omega_j}{AUP} \]

**Linear Expressions for Mixing Ternary Clusters**

Binary:

\[ C_{mix}^{(1-2)} = \sum_{j=1}^{3} \beta_j \cdot C_{j} \]

**Feasibility Conditions**

1. **Necessary Condition**
   Match clustering target.

2. **Sufficient Condition**
   Match AUP value of sink.

3. **Property Constraints**
   \[ p_{j,\text{min}} \leq p_j \leq p_{j,\text{max}} \]

**Components:**

- \( x_1, x_2, x_3 \)

• **Benefits**
  - Provides the ability to link different scales.
  - Can handle combinatorial intensive problems.
  - Universal in application.
**Example 1: Process Design**

**Stream Characterization**
1. Objectionable Material (OM)
2. Absorption coefficient (k)
3. Reflectivity ($R_\infty$)

**Given Information**
1. Property and flow data for fibers and broke
2. Paper machine feed condition constraints
### Minimum Fiber Usage

Assuming full recycle and interception capabilities, the minimum fiber usage will be **70 ton/hr**

### Solution Strategies

1. Direct recycle
2. Interception and recycle

### Table

<table>
<thead>
<tr>
<th>Property</th>
<th>Operator</th>
<th>Fibers</th>
<th>Broke</th>
<th>Paper machine</th>
</tr>
</thead>
<tbody>
<tr>
<td>OM (mass fraction)</td>
<td>OM</td>
<td>0.000</td>
<td>0.115</td>
<td>0.00 – 0.02</td>
</tr>
<tr>
<td>$k$ (m$^2$/g)</td>
<td>$k$</td>
<td>0.0012</td>
<td>0.0013</td>
<td>0.00115 – 0.00125</td>
</tr>
<tr>
<td>$R_\infty$</td>
<td>$(R_\infty)^{5.92}$</td>
<td>0.82</td>
<td>0.90</td>
<td>0.80 – 0.90</td>
</tr>
<tr>
<td>Flowrate (ton/hr)</td>
<td>------</td>
<td><strong>100</strong></td>
<td><strong>30</strong></td>
<td><strong>100 – 105</strong></td>
</tr>
</tbody>
</table>
Optimal Mixing
Minimum fiber arm, corresponding to flowrate of 83 ton/hr

Reverse Problem
Flowrates unknown, i.e. mixing point yields flow fractions
Reverse Problem
70 ton/hr fiber usage, with current feed mixture point, identify new cluster point for broke.

Intercepted Broke
OM = 0.067
k = 0.0011
$R_\infty = 0.879$
Molecular Clusters 1:2

**Process Property Operators**

\[
\psi_j^P = \sum_{s=1}^{N_s} x_s \cdot P_{js}
\]

Linear Expression for Mixing 2 Ternary Clusters

\[
C_j^\text{MIX} = \sum_{s=1}^{N_s} \beta_s \cdot C_{js}
\]

**Molecular Property Operators**

\[
\psi_j^M = \sum_{g=1}^{N_g} n_g \cdot P_{jg}
\]

G_1 and G_2 are added linearly on the ternary diagram. The location of \( \beta_1 \) corresponds to the location of G_1-G_2 molecule

\[
AUP = \sum_{j=1}^{N_j} \Omega_j
\]

\[
C_j = \frac{\Omega_j}{AUP}
\]

\[
\beta_1 = \frac{n_1 \cdot AUP_1}{n_1 \cdot AUP_1 + n_2 \cdot AUP_2}
\]
$\beta_1$, the visualization arm, corresponds to the location of G1-G2 molecule

$$\beta_1 = \frac{n_1 \cdot AUP_1}{n_1 \cdot AUP_1 + n_2 \cdot AUP_2}$$

**Feasibility**

**Necessary Conditions**
1. Free bond number is zero.
2. Match clustering target
3. Match AUP range of sink

**Sufficient Condition**
Check property value with sink including Non-GC properties
Example 2: Molecular Synthesis

- **Blanket Wash Solvent Design**
  - Solved as MINLP by Sinha and Achenie (2001)

- **Problem Statement**
  - Design blanket wash solvent for phenolic resin printing ink
  - Molecules designed from 7 possible groups, with a max. chain length of 7 groups

<table>
<thead>
<tr>
<th>Property</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hv (kJ/mol)</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>Tb (K)</td>
<td>350</td>
<td>400</td>
</tr>
<tr>
<td>Tm (K)</td>
<td>150</td>
<td>250</td>
</tr>
<tr>
<td>VP (mmHg)</td>
<td>100</td>
<td>---</td>
</tr>
<tr>
<td>R^i</td>
<td>0</td>
<td>19.8</td>
</tr>
</tbody>
</table>
Blanket Wash Solvent 1:7

- Visualization limits problem to three properties
- Heat of vaporization, boiling and melting temperatures are used, with vapor pressure and solubility used as final screening properties

**Property Prediction (GCM)**

\[
\Delta H_v - h_{vo} = \sum_i g_i \cdot h_{vi}
\]

\[
T_b = t_{bo} \cdot \ln \sum_i g_i \cdot t_{b1_i}
\]

\[
T_m = t_{mo} \cdot \ln \sum_i g_i \cdot t_{m1_i}
\]

**Molecular Property Operators**

\[
\Delta H_v - h_{vo} = \sum_i n_g \cdot h_{v1} \quad , \psi_{ref} = 20
\]

\[
\exp \left( \frac{T_b}{t_{bo}} \right) = \sum_{g=1}^{N_g} n_g \cdot t_{b1} \quad , \psi_{ref} = 100
\]

\[
\exp \left( \frac{T_m}{t_{mo}} \right) = \sum_{g=1}^{N_g} n_g \cdot t_{m1} \quad , \psi_{ref} = 7
\]
Blanket Wash Solvent 3:7

Molecular Groups
- G1: CH₃
- G2: CH₂
- G3: CH₂O
- G4: CH₃O
- G5: CH₂CO
- G6: CH₃CO
- G7: COOH
Blanket Wash Solvent 4:7

Candidate Molecules
M1 : CH₂O-CH₂-(CH₂O)₂
M2 : CH₃-CH₂-CH₂O-CH₃O
M3 : CH₃-CH₂-(CH₂O)₂-CH₃
M4 : CH₃O-(CH₂)₃-CH₃
M5 : CH₃O-CH₂O-CH₃O
M6 : CH₂O-(CH₂O)₂-CH₂O
M7 : CH₃-CH₂CO-CH₃
M8 : CH₃-(CH₂)₅-CH₃
M9 : CH₃CO-COOH
M10 : CH₃CO-(CH₂)₂-COOH
M11 : CH₃-CH₂-(CH₂O)₂
**Feasible formulations from Visual Synthesis**

<table>
<thead>
<tr>
<th>Formulations</th>
<th>AUP</th>
<th>(H_v) kJ/mol</th>
<th>(T_b) K</th>
<th>(T_m) K</th>
<th>VP mmHg</th>
<th>(R^i) MPa</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>3.20</td>
<td>33.91</td>
<td>359.14</td>
<td>201.24</td>
<td>1117.85</td>
<td>10.88</td>
</tr>
<tr>
<td>M2</td>
<td>3.08</td>
<td>33.99</td>
<td>355.34</td>
<td>189.86</td>
<td>1240.95</td>
<td>15.39</td>
</tr>
<tr>
<td>M3</td>
<td>3.10</td>
<td>34.67</td>
<td>364.49</td>
<td>183.38</td>
<td>963.57</td>
<td>12.83</td>
</tr>
<tr>
<td>M4</td>
<td>3.17</td>
<td>35.81</td>
<td>363.09</td>
<td>186.26</td>
<td>1001.85</td>
<td>18.09</td>
</tr>
<tr>
<td>M5</td>
<td>3.47</td>
<td>36.15</td>
<td>370.61</td>
<td>211.95</td>
<td>811.54</td>
<td>15.82</td>
</tr>
<tr>
<td>M6</td>
<td>3.61</td>
<td>36.74</td>
<td>382.51</td>
<td>216.49</td>
<td>578.05</td>
<td>11.31</td>
</tr>
<tr>
<td>M7</td>
<td>3.17</td>
<td>35.10</td>
<td>354.80</td>
<td>193.13</td>
<td>1259.28</td>
<td>16.84</td>
</tr>
<tr>
<td>M8</td>
<td>3.28</td>
<td>38.31</td>
<td>379.07</td>
<td>175.55</td>
<td>638.03</td>
<td>19.77</td>
</tr>
<tr>
<td>M9</td>
<td>6.79</td>
<td>68.87</td>
<td>457.92</td>
<td>286.76</td>
<td>56.69</td>
<td>13.83</td>
</tr>
<tr>
<td>M10</td>
<td>7.79</td>
<td>78.17</td>
<td>494.54</td>
<td>297.68</td>
<td>16.55</td>
<td>12.68</td>
</tr>
<tr>
<td>M11</td>
<td>7.83</td>
<td>74.75</td>
<td>535.55</td>
<td>292.08</td>
<td>3.86</td>
<td>11.33</td>
</tr>
</tbody>
</table>

**Application of feasibility conditions**

- All formulations satisfy the first two necessary conditions
- M9-M11 fail to satisfy the AUP range of the sink
## Blanket Wash Solvent 6:7

### Feasible formulations from Visual Synthesis

<table>
<thead>
<tr>
<th>Formulations</th>
<th>AUP</th>
<th>$H_v$ (kJ/mol)</th>
<th>$T_b$ (K)</th>
<th>$T_m$ (K)</th>
<th>VP (mmHg)</th>
<th>$R^{ij}$ (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>3.20</td>
<td>33.91</td>
<td>359.14</td>
<td>201.24</td>
<td>1117.85</td>
<td>10.88</td>
</tr>
<tr>
<td>M2</td>
<td>3.08</td>
<td>33.99</td>
<td>355.34</td>
<td>189.86</td>
<td>1240.95</td>
<td>15.39</td>
</tr>
<tr>
<td>M3</td>
<td>3.10</td>
<td>34.67</td>
<td>364.49</td>
<td>183.38</td>
<td>963.57</td>
<td>12.83</td>
</tr>
<tr>
<td>M4</td>
<td>3.17</td>
<td>35.81</td>
<td>363.09</td>
<td>186.26</td>
<td>1001.85</td>
<td>18.09</td>
</tr>
<tr>
<td>M5</td>
<td>3.47</td>
<td>36.15</td>
<td>370.61</td>
<td>211.95</td>
<td>811.54</td>
<td>15.82</td>
</tr>
<tr>
<td>M6</td>
<td>3.61</td>
<td>36.74</td>
<td>382.51</td>
<td>216.49</td>
<td>578.05</td>
<td>11.31</td>
</tr>
<tr>
<td>M7</td>
<td>3.17</td>
<td>35.10</td>
<td>354.80</td>
<td>193.13</td>
<td>1259.28</td>
<td>16.84</td>
</tr>
<tr>
<td>M8</td>
<td>3.28</td>
<td>38.31</td>
<td>379.07</td>
<td>175.55</td>
<td>638.03</td>
<td>19.77</td>
</tr>
<tr>
<td>M9</td>
<td>6.79</td>
<td>68.87</td>
<td>457.92</td>
<td>286.76</td>
<td>56.69</td>
<td>13.83</td>
</tr>
<tr>
<td>M10</td>
<td>7.79</td>
<td>78.17</td>
<td>494.54</td>
<td>297.68</td>
<td>16.55</td>
<td>12.68</td>
</tr>
<tr>
<td>M11</td>
<td>7.83</td>
<td>74.75</td>
<td>535.55</td>
<td>292.08</td>
<td>3.86</td>
<td>11.33</td>
</tr>
</tbody>
</table>

### Application of feasibility conditions

- Checking property values with sink including Non-GC properties (VP, solubility), the sufficient conditions are satisfied for remaining formulations
Candidate molecules M1-M7 identified visually by the developed method correspond to solutions found by the MINLP approach used by Sinha and Achenie (2001).

Although valid formulation, heptane (M8) is flammable hence not an ideal solvent.

Valid Formulations:

- M1: CH₃O-CH₂-(CH₂O)₂
- M2: CH₃-CH₂-CH₂O-CH₃O
- M3: CH₃-CH₂-(CH₂O)₂-CH₃
- M4: CH₃O-(CH₂)₃-CH₃
- M5: CH₃O-CH₂O-CH₃O
- M6: CH₂O-(CH₂O)₂-CH₂O
- M7: CH₃-CH₂CO-CH₃
- M8: CH₃-(CH₂)₅-CH₃

Cyclical compounds:

Ethers:

MEK
Integrated Design Approach

Stream Properties & Unit Constraints

Clusters

Process Design

Property Targets

Clusters

Molecular Formulations

Molecular Design
Objective
To maximize the use of off-gas condensate and to minimize fresh solvent use to the degrease.

Stream Characterization
- Sulfur Content (S)
- Molar Volume (Vm)
- Vapor Pressure (VP)

Diagram:
- Absorber
  - 4.4 kg/min Fresh Solvent 2
  - Spent Organics for incineration
- Solvent Regeneration
  - Recycle Solvent
- Degreasing
  - Metal
  - Evaporated Solvent Off-gas to Flare or Condensation for Reuse
- Finish Processing
  - Fresh Solvent 1 36.6 kg/min
  - Degreased Metal
• **Degreaser Feed Constraints**

<table>
<thead>
<tr>
<th>Property</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>S (%)</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$V_m$ (cm$^3$/mol)</td>
<td>90.09</td>
<td>487.80</td>
</tr>
<tr>
<td>VP (mmHg)</td>
<td>1596</td>
<td>3040</td>
</tr>
</tbody>
</table>

• **Property Operator Mixing Rules**

\[
S_M = \sum_{s=1}^{N_s} x_s \cdot S_s
\]

, $S^{\text{ref}} = 0.5 \text{ wt\%}$

\[
V_m = \sum_{s=1}^{N_s} x_s \cdot V_{m_s}
\]

, $V_{m}^{\text{ref}} = 80 \text{ cm}^3/\text{mol}$

\[
VP_M^{1.44} = \sum_{s=1}^{N_s} x_s \cdot VP_s^{1.44}
\]

, $VP^{\text{ref}} = 760 \text{ mmHg}$
**Visualization of Process Design Problem**

**VOC Condensation Data**

Sulfur content, density and vapor pressure data given for temperature range 480K-515K.

Metals Degreasing 2:9
**Visualization of Process Design Problem**

- **Point A & B** dictate property constraint targets

**Conditions**

- Condenser operates @ 500 K
- Feed Solvent must have zero sulfur content
**Values from Process Design Visual Solution**

<table>
<thead>
<tr>
<th>S (%)</th>
<th>V&lt;sub&gt;m&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;/mol)</th>
<th>VP (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>102.09</td>
<td>1825.37</td>
</tr>
<tr>
<td>0</td>
<td>720.75</td>
<td>3878.66</td>
</tr>
</tbody>
</table>

**Molecular Property Constraints**

<table>
<thead>
<tr>
<th>Hv (kJ/mol)</th>
<th>V&lt;sub&gt;m&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;/mol)</th>
<th>VP (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>102.09</td>
<td>1825.37</td>
</tr>
<tr>
<td>100</td>
<td>720.75</td>
<td>3878.66</td>
</tr>
</tbody>
</table>
Metal Degreasing 5:9

**Property Prediction (GCM)**

\[
\Delta H_v = h_{vo} + \sum_i n_i \cdot h_{vl_i}
\]

\[
V_m = d + \sum_i n_i \cdot v_{li}
\]

\[
T_{bp} = t_{bo} \cdot \ln \sum_i n_i \cdot t_{bl_i}
\]

\[
\log_{10} VP = 5.58 - 2.7 \left( \frac{T_{bp}}{T} \right)^{1.7}
\]

**Molecular Property Operators**

\[
\Delta H_v - h_{vo} = \sum_{g=1}^{N_g} n_g \cdot h_{vl_g}, \psi^{ref} = 20
\]

\[
V_m - d = \sum_{g=1}^{N_g} n_g \cdot v_{l_g}, \psi^{ref} = 100
\]

\[
\exp \left( \frac{T_{bo}}{t_{bo}} \right) = \sum_{g=1}^{N_g} n_g \cdot t_{bl_g}, \psi^{ref} = 7
\]

Non-GC Property
Molecular Fragments

G1: CH₃
G2: CH₂
G3: CH₂O
G4: CH₂N
G5: CH₃N
G6: CH₃CO
G7: COOH

Visualization of Molecular Design Problem
Visualization of Molecular Design Problem

Candidate Molecules

M1 CH₃-(CH₂)₅-CH₃CO
M2 CH₃CO-(CH₂)₂-CH₃CO
M3 (CH₃)₃-(CH₂)₅-CH₂N
M4 CH₃-(CH₂)₂-COOH
M5 (CH₃)₂-CH₃CO-CCL
M6 -(CH₂O)₅- ring
M7 CH₃-(CH₂)₂-CH₃N-COOH
• **Formulations from Visual Design**

<table>
<thead>
<tr>
<th>Formulation</th>
<th>AUP</th>
<th>Tb (K)</th>
<th>Hv  (kJ/mol)</th>
<th>Vm (cm³/mol)</th>
<th>VP (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>5.06</td>
<td>450.58</td>
<td>53.19</td>
<td>156.85</td>
<td>2078.98</td>
</tr>
<tr>
<td>M2</td>
<td>4.71</td>
<td>448.54</td>
<td>54.13</td>
<td>118.03</td>
<td>2163.90</td>
</tr>
<tr>
<td>M3</td>
<td>5.11</td>
<td>407.29</td>
<td>40.95</td>
<td>109.41</td>
<td>2002.07</td>
</tr>
<tr>
<td>M4</td>
<td>4.86</td>
<td>438.97</td>
<td>63.29</td>
<td>93.39</td>
<td>2606.12</td>
</tr>
<tr>
<td>M5</td>
<td>4.02</td>
<td>419.20</td>
<td>49.88</td>
<td>121.14</td>
<td>4241.48</td>
</tr>
<tr>
<td>M6</td>
<td>4.19</td>
<td>438.11</td>
<td>44.22</td>
<td>137.66</td>
<td>3308.12</td>
</tr>
<tr>
<td>M7</td>
<td>5.71</td>
<td>485.01</td>
<td>70.24</td>
<td>112.32</td>
<td>1037.99</td>
</tr>
</tbody>
</table>

• **Application of Feasibility Conditions**
  
  – All formulations satisfy first two necessary conditions
  
  – M5 and M6 fail to satisfy sink AUP range
  
  – M3 and M7 did not match Non-GC property value
  
  – M1, M2 and M4 are valid solvent candidates
Metal Degreasing  9:9

Visualization of Process Design Solution

19.36 kg/min of 2,5-hexadione as fresh solvent
Product Design Challenges


- **Strategy Execution**
  - “The power of 21st century computing is making it possible to **predict certain structures and properties from fundamental principles**”
  - “One of the important lessons learned...is the profound importance of experimental results to calibrate and validate computational methods and fill gaps in theoretical understanding...best practiced with **complimentary experimental and theoretical approaches**”
  - “The properties of materials are controlled by a multitude of separate and often competing mechanisms that operate over a wide range of length and time scales...**the linkage of various methods remains a great challenge**.”
Way Forward

- Two paths: (1) “Systematic reduction of the number of alternatives through heuristics” and (2) “optimization of the set of all potential alternatives through mathematical programming”

- “Requisite properties will be performance attributes rather than fundamental physico-chemical, or bio-chemical properties”

- “New property estimation techniques will need to be identified and proven...since formulated products are often multi-phase structured products”

- Methodology will “not eliminate the need for product experimentation in the laboratory, but serve to guide and focus product experimentation”

Cryo-SEM micrographs of a lamellar-structured hair conditioner manufactured under low deformation rates (top) and high deformation rates (bottom) (Edwards, 1998).

• **Multiscale Systems**
  - Integration of various product and process relationships so as to ensure complete solutions and global optimums
  - Utilize reverse problem formulations to generate multiple match solutions
Motivations and Challenges

Simultaneous Matching of All Scales?

Fundamental Principles of Chemical Engineering

Properties

Continuous Decisions (e.g. operating conditions)

Integrating design capable of delivering desired performance

Consumer Attributes (Macro/Mega Scale)

Chemometric Models

Microstructure (Meso Scale)

Coarse-Graining and Chemometric Models

Molecular Design

Discrete Decisions

CAPD / CAMD (Molecular Scale)

Group Contribution and Chemometric Models

Given a set of molecular groups to be screened (building blocks)

Ab Initio Methods (Quantum Scale)

QM, DFT Models

Atomistic Methods (Atomic Scale)

MM, MD Simulations

Desired process (e.g. product specifications)

Constraints on properties: values obtained by experimental measurements, targeting optimisation for performance and sustainability

Designed Molecules (raw materials, MSA’s)
Motivations and Challenges

- **Research Focus**
  - Linkage of “Complimentary Experimental and Theoretical Approaches”
    - Chemometric Inference Models
      - Mixture Design of Experiments (MDOE)
      - Decomposition Techniques (PCA)
      - Stochastic Simulations (Monte Carlo)
    - Semi-Empirical Models
      - Higher Order Group Contribution
      - Connectivity Indices
      - Molecular Graph Theory

Path 1
Path 2
• **Objective**
  - Use a reverse problem formulation in the appropriate linearized subspace

• **Method Development**
  - Method includes
    1. Development of a centralized framework
    2. Development of an algorithm that chooses the appropriate properties to bridge the length scales
    3. Use of a property clustering algorithm to load descriptions into the framework and visualize the solution
**Empirical Method**
- Utilizes direct Attribute-Component relationship
- Limited to selection of components only; each new molecule requires additional experimentation
- Limited to linear correlations for global optimization
- Subject to combinatorial explosion when developing model parameters

**Semi-Empirical Method**
- Utilizes both an Attribute-Property relationship and a Property-Component relationship
- Enumerates all candidate molecules using molecular design techniques based on group-based indices
- Uses linear or nonlinear correlations of attributes for global optimization
- Less subjective to combinatorial explosion when developing model parameters

In addition, both methods also utilize Property Clustering and Chemometric Techniques
• **Chemometrics**
  - “Chemometrics is the science of relating measurements made on a system or process to the state of the system or process via application of mathematical or statistical methods” – *The International Chemometrics Society (ICS)*
  - “There are techniques for collecting good data (optimization of experimental parameters, design of experiments, calibration, signal processing) and for getting information from these data (statistics, pattern recognition, modeling, structure-property-relationship estimations)” – *Wikipedia*

• **Focus**
  - Design of Experiments (MDOE, etc.)
  - Characterization Techniques (Spectroscopy, etc.)
  - Decomposition Techniques (PCR, etc.)
Mixture Design

What is Mixture Design?
- Mixture Design is a Design of Experiments (DOE) tool used to determine the optimum combination of chemical constituents that deliver a desired response using a minimum number of experimental runs.

Experimentation
\[ Y = f(X) \]

Interpretation /
Model Selection

Prediction

<table>
<thead>
<tr>
<th>Trial</th>
<th>( X_1 )</th>
<th>( X_2 )</th>
<th>( X_3 )</th>
<th>( \hat{y} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.68</td>
<td>0.16</td>
<td>0.16</td>
<td>15.2</td>
</tr>
<tr>
<td>2</td>
<td>0.16</td>
<td>0.68</td>
<td>0.16</td>
<td>12.2</td>
</tr>
<tr>
<td>3</td>
<td>0.16</td>
<td>0.16</td>
<td>0.68</td>
<td>15.2</td>
</tr>
<tr>
<td>4</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>14.6</td>
</tr>
</tbody>
</table>

\[ y = b_0 + \sum_{i=1}^{q} b_i x_i + \varepsilon \]

\[ B = X'Y(X'X)^{-1} \]

\[ \hat{y} = \sum_{i=1}^{q} b_i^* x_i + \varepsilon \]

\[ Y = B^* X \]

\[ Y = BX \]
• **Mixture Design Limitations**
  – Suffers from combinatorial problems
    • 7 components = 25 independent ternary plots per property
  – Evaluation of multiple effects is difficult

• **New method of visualizing mixture designs**
  – Must handle combinatorial intensive problems
  – Must be easy to visualize
  – Must be universal in its application
MDOE Using Property Clusters

Conventional MDOE: 7 Components & 3 Properties = 75 plots

Property Clustering MDOE: 7 Components & 3 Properties = 1 plot
Example 4: Mixture Design

- **Polymer Blend Study**
  - Optimization of a polymer blend of spun yarn for use in rope for modern racing sailboats

- **Objective**
  - Optimize a ternary or smaller polymer blend to deliver the specified product attributes of high strength, low stretch, and high floatability
Polymer Blend Design 1:2

- **Attribute – Property Relationship**
  2. Stretch – Thread Elongation (Cornell, 2002)
  3. Floatability – Specific Volume (Eden et. al, 2003)

- **Property Targets – Feasibility Range**

<table>
<thead>
<tr>
<th>Property</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thread Elongation (kgf)</td>
<td>12.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Knot Strength (lbf)</td>
<td>12.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Specific Volume (cm³/g)</td>
<td>1.063</td>
<td>1.111</td>
</tr>
</tbody>
</table>

- **Polymer Candidates**
  1. LDPE
  2. Polystyrene
  3. Polypropylene
  4. Nylon 6,6
Polymers Blend Design 2:2

MDOE

Thread Elongation

Knot Strength

Specific Volume

Polystyrene has the largest effect, Polypropylene should be used as filler.
Case Study: Excipient Design

- **Acetaminophen Tablet Design**
  - “Optimization of poorly compactable drug tablets manufactured by direct compression using mixture experimental design” – Martinello *et al.*, 2006

- **Objective:**
  - Determine if a ternary or smaller mixture can deliver a directly compressible compact of acetaminophen
Case Study: Excipient Design

- **Powder Properties**
  - Angle of Repose (RA)
  - Compressibility (C)
  - Water Content (W)

- **Candidates**

<table>
<thead>
<tr>
<th>Excipient</th>
<th>Industry Name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Microcel 102</td>
<td>Filler-Binder</td>
</tr>
<tr>
<td>2</td>
<td>Kollidon VA 64</td>
<td>Binder</td>
</tr>
<tr>
<td>3</td>
<td>Flowlac</td>
<td>Filler-Binder</td>
</tr>
<tr>
<td>4</td>
<td>Kollidon CL 30</td>
<td>Disintegrant</td>
</tr>
<tr>
<td>5</td>
<td>PEG 4000</td>
<td>Glidant-Lubricant</td>
</tr>
<tr>
<td>6</td>
<td>Aerosil 200 Pharma</td>
<td>Glidant-Lubricant</td>
</tr>
<tr>
<td>7</td>
<td>Magnesium Stearate</td>
<td>Lubricant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Property</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA (deg)</td>
<td>0.0</td>
<td>30.0</td>
</tr>
<tr>
<td>C (%)</td>
<td>0.0</td>
<td>32.0</td>
</tr>
<tr>
<td>W (wt.%)</td>
<td>1.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>

\[
RA_M = \sum_{s=1}^{Ns} x_s \cdot b_s^*(RA_s)
\]
\[
C_M = \sum_{s=1}^{Ns} x_s \cdot b_s^*(C_s)
\]
\[
W_M = \sum_{s=1}^{Ns} x_s \cdot b_s^*(W_s)
\]
• **Constraint Check**
  - Distance to real cluster space indicative of **STRONG** effect
  - All formulations meet all powder and tablet constraints independently

**Possible Formulations**

- 2-6
- 4-6
- 1-2-6
- 1-3-6
- 1-4-6
- 1-5-6
- 2-3-6
- 2-4-6
- 2-5-6
- 2-7-6
- 3-4-6
- 4-5-6
- 4-7-6
Attribute-Component Conclusions

- Handles combinatorial explosion
- Representation of combined effects of each component
- Requires the use of negative cluster space handled by constraint $AUP > 0$

Challenges

- What about “new” molecules?
- How to deal with secondary colinearity or nonlinearity?
- What about “principal” properties?

One Answer: Use Subspace Mapping w/ Latent Variable Models
• **Principal Property Method**
  - DOE/PLS used to determine attribute-property relationship
  - PCA to find principal properties
  - Linearized principal property-component relationships are used to determine structure
  - Molecular design via a modified group contribution

\[
Q_{\text{mix}} = RQ \\
Q_{\text{mix}} = N_gQ
\]
Summary

• **Property Clusters**
  – Provide a framework to solve property driven processes without any commitment to components.
  – Unified framework for simultaneous solution of process & molecular design problems.
  – Enables visualization and reduces problem dimensionality.
  – Maps attribute data down to subspace with orthogonal linear functions, thus ensuring global optimums and complete candidate sets are found.
Future Directions

• Property Clusters
  – Expand property description options through molecular signature and characterization techniques
  – Integrate topological indices with group contribution based flowsheet design
  – Include microstructure information via crystallinity algorithms, both deterministic and stochastic
  – Analyze the proper information pathways through ontology, network component analysis, and other algorithms
Acknowledgements

- **Funding**
  - NSF CAREER Program
  - DOE Consortium for Fossil Fuel Science (CFFS)

- **Students**
  - Charles Solvason & Nishanth Chemmangattuvalappil
Dr. Mario Richard Eden
Mary & John H. Sanders Associate Professor
Department of Chemical Engineering
222 Ross Hall
Auburn University, AL 36849-5127

Phone: (334) 844-2064
Fax: (334) 844-2063
Email: edenmar@auburn.edu
Web: www.eng.auburn.edu/users/edenmar