Model Integration Using Ontology
Input-Output Matching

Linsey Koo, Franjo Cecelja

Process and Information Systems Engineering Research Centre
Chemical and Process System Engineering
University of Surrey, Guildford, United Kingdom

Overview

- Challenges in Biorefining Modelling
- Existing Framework
- Ontological Approach of Model Integration
- Demonstration of Input-Output Matching
- Conclusion & Future work
Challenges of biorefining modelling

- **Complexity** of biorefining modelling & modelling methods
  - Require a wide range of expertise and time

- **Independently** developed models
  - Different users
  - Best suited modelling tools

- **Various** ways of developing models
  - Achieve various aims and purposes
  - Cause inconsistencies of the models

- **Implicit** models
  - Lack of awareness of existing models
  - Limits the potential of reusability of these models

Existing Framework: CAPE-OPEN

- Address the question of standardisation of interfaces
  - Enable interoperability between simulator software components

- Framework built around a middleware
  - Computation of physical properties
  - Simulation of particular unit operation
  - Numerical solutions for specific mathematical problems

Disadvantage: Lack of flexibility

Ontological Approach

- Build upon the existing CAPE-OPEN framework
  - Replaces the CAPE-OPEN middleware with a more intuitive and flexible semantic repository
  - Semantic repository resides on web
- Increase awareness of existing models
  - Increase model reusability and integration
  - Reduce the developing time, efforts, and the need for expertise.

Domain Ontology

- Describe models representing biorefining processes
  - Support explicit description of model and data & enhance consistency
Domain Ontology

\[ O = \{R^I, H_C, R^C, R^E, f_E, S_I^I\} \]

\[ f_R(D) = \text{dom} R^I \rightarrow f_D: S_I^I \rightarrow S_D^D \]

\[ f_R(S) = \text{ran}g R^I \rightarrow f_S: S_I^I \rightarrow S_S^S \]

Hence binary relationship:

- \( S_D^D \) and \( S_S^S \) universal and existential quantifiers over properties \( R^I \) of \( S_I^I \),
- \( S_D^D \) and \( n, n \in \mathbb{N} \), cardinality quantifiers over properties \( R^D \) of \( S_I^I \),
- \( S_S^S \) and \( v, v \in S_I \cup \mathbb{N} \), equality quantifiers over properties \( R^E \) of \( S_I^I \).

Semantic Model Description

- Semantic Web Services (SWS) ontology
  - Describe model/data defining biorefining
  - ServiceProfile: what it does (outputs, functionality)
  - ServiceGrounding: environment (software, preconditions)
  - ServiceModel: requirements of the models (key parameters)
Semantic Model Description

- **Semantic Web Services** (SWS) ontology
  - Describe model/data defining biorefining
  - ServiceProfile: what it does (outputs, functionality)
  - ServiceGrounding: environment (software, preconditions)
  - ServiceModel: requirements of the models (key parameters)

---

Ontological Approach

- Process of model integration
  - Register as an instance of the domain ontology (Semantic Web Services)
  - Publish in the purposely built public repository for I-O matching

- Use **Ontology Web Service Description** (OWL-S) framework
  - Discovery stage
  - Selection stage
  - Composition stage
  - Execution stage
Partial Matching

- Facilitate the **flexibility** of model integration through I/O partial matching
  - Strong synthesis capabilities and functions
  - Invite **degrees of freedom**

- **Methods**
  - Semantic Matching
  - Property Matching

Semantic Matching

- Measure **semantic similarities**: Input/Output types & properties
  - Facilitate the **flexibility** of model integration
  - Partially satisfying the input criteria
**Property Matching**

-距离测量方法: **数值**
  - 向量比较
  - 玄同相似度
  - 欧几里德相似度
  - 聚合相似度

\[
h_{k}^{v,c} = \frac{P_{\text{r}} \cdot P_{\text{i}}}{\|P_{\text{r}}\| \cdot \|P_{\text{i}}\|} = \frac{\sum_{i=1}^{n} P_{r,i} \times P_{i,i}}{\sqrt{\sum_{i=1}^{n} (P_{r,i})^2} \times \sqrt{\sum_{i=1}^{n} (P_{i,i})^2}}
\]

\[
h_{k}^{v,e} = 1 - \frac{\sqrt{\sum_{i=1}^{n} (P_{r,i} - P_{i,i})^2}}{\max \sqrt{\sum_{i=1}^{n} (P_{r,i} - P_{i,i})^2}}
\]

\[
h_{k}^{v} = \frac{h_{k}^{v,c} + h_{k}^{v,e}}{2}
\]


**Backward Matching Process**

- 开发基于输入-输出 (I-O) 匹配的技术
  - 高效而准确的模型和数据发现
  - 避免过程的扩展，与前向匹配对比

**Forward Matching Process**
(资源作为请求者)

**Backward Matching Process**
(解决方案作为请求者)
### Co-Fermentation (Requestor)
- using the bacterium Zymomonas Mobilis
- Produce 315M litre of ethanol per year from corn stover
- Operating temperature **41 degree C**
- **Resident time** of 1.5 days

### Pretreatment
- **Pretreatment Ontology**

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilute Acid and Enzymatic Hydrolysis (DAEH)</td>
<td></td>
</tr>
<tr>
<td>Input: Barley Straw</td>
<td>Output: 51.6% sugar @ 90°C</td>
</tr>
<tr>
<td>Ammonia Fibre Expansion (AFEX)</td>
<td>Output: Switchy Jesse @ 23°C</td>
</tr>
</tbody>
</table>

### Demonstration: Semantic Measure

- **Semantic Similarity**
  - Material Type

![Semantic Measure Diagram]

60% 20%
Demonstration: Property Measure

- Property Similarity Measure
  - Physical Property
  - Quantity, Temperature, Pressure, Composition
- Registration
  - Collect **explicit** information by parsing the ontology

### Technology Type: Pre-treatment

<table>
<thead>
<tr>
<th>Technology Type</th>
<th>Model Function</th>
<th>Model Output</th>
<th>Model Input</th>
<th>Output Quantity</th>
<th>Operating Temperature</th>
<th>Input Composition</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilute Acid and Enzymatic Hydrolysis (DAEH)</td>
<td>Sugar</td>
<td>Switchgrass</td>
<td>46 t/h</td>
<td>50ºC</td>
<td>34.2% Glucan, 22.1% Xylan, 3.1% Arabinan</td>
<td>95% Glucose yield, 93% Xylose yield</td>
<td></td>
</tr>
</tbody>
</table>

### Technology Type: Conversion Reaction

<table>
<thead>
<tr>
<th>Technology Type</th>
<th>Model Function</th>
<th>Model Output</th>
<th>Model Input</th>
<th>Operating Temperature</th>
<th>Input Composition (CornStover)</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Fermentation</td>
<td>Ethanol</td>
<td>Sugar from CornStover</td>
<td>56 t/h</td>
<td>41ºC</td>
<td>37.4% Glucan, 22.1% Xylan</td>
<td>99.7% Glucose yield, 94.3% Xylose yield</td>
</tr>
</tbody>
</table>

### Aggregated Similarity

- Semantic Similarity: 60%
- Property Similarity: 63%

### Cohesive Matching

- Result

<table>
<thead>
<tr>
<th>Technology Type</th>
<th>Mode Function</th>
<th>Model Output</th>
<th>Model Input</th>
<th>Output Quantity</th>
<th>Operating Temperature</th>
<th>Input Composition</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Fermentation</td>
<td>Ethanol</td>
<td>Sugar from CornStover</td>
<td>56 t/h</td>
<td>41ºC</td>
<td>37.4% Glucan, 22.1% Xylan</td>
<td>99.7% Glucose yield, 94.3% Xylose yield</td>
<td></td>
</tr>
</tbody>
</table>

### Aggregated Similarity

- Semantic Similarity: 20%
- Property Similarity: 50%
Conclusions

- **Semantic Web Service descriptions** for conversion technologies and resources are established to enable Input-Output matching

- Model Integration through **Input-Output matching** was implemented in OWL-S Framework

- **Backward matching process** for model discovery

- Facilitate the **flexibility** of model integration using partial matching

- Research continues towards identifying **key parameters** and understanding of the **impact of partial matching**

Acknowledgement