An Objective-based Approach for Semantic Validation of Emergence in Component-based Simulation Models

Claudia Szabo* and Yong Meng Teo+
Department of Computer Science
National University of Singapore
email: teoym@comp.nus.edu.sg

*Department of Computer Science
University of Adelaide, Australia

+Visiting Professor
Shanghai Advanced Research Institute
Chinese Academy of Sciences
Motivation

“The whole is greater than the sum of its parts” - Aristotle
Outline

• Introduction
• Related Work
• Objective-based Validation Approach
• Example: Flock of Birds
• Conclusion
Complex Systems, Simulation & Emergence

• Ubiquitous in today’s systems
  – Single component => many components
  – Simple components => large number of states and attributes

• Simulation
  – Increased flexibility
  – A system is complex if it exhibits one or more properties not resulting from the properties of its interconnected parts.

• Emergence: malign and benign effects
  – Ethernet capture effect [20]
  – router synchronization problems [9]
  – load-balance failures in multi-tiered distributed systems [17]
  – herding behavior among clients
  – flocking of birds, ant colonies, crowd behaviors
  – ...

21 August 2014  PADS, Zhangjiajie, China, July 2012
Complex Systems Research

Complexity Science

- Theories
  - Network theory
    - Continuous dynamics
  - Dynamical systems theory
    - Discrete dynamics
- Statistical mechanics
- Biology
- Traffic systems
- Generic algorithms
- Swarm algorithms

Complex systems design & engineering

- Domain-specific
  - Component-based modeling & simulation
    - Cellular automata
    - Agent-based model

- Algorithms
- Tools & frameworks

Domain-specific

Cellular automata

Tools & frameworks

Entropy

Continuous dynamics

Discrete dynamics

Granger causality
Modeling and Simulation

Traditional Approach

Simulation Problem \( P \) \( \rightarrow \) Model(s) \( M \) \( \rightarrow \) Implementation(s) \( \rightarrow \) Execution

Validation

Verification

Semantic Composability Validation

Syntactic Composability Verification

Component-based Approach

PADS, Zhangjiajie, China, July 2012
Previous Work

Crosscutting Issues in Component-based Model Development

- Syntactic Composability
- Model Discovery
- Model Selection
- Semantic Composability

- Component representation
- Model reuse
- Data encapsulation
- Loose coupling
- Component repository
- Heterogeneous components
- Domain knowledge

21 August 2014
PADS, Zhangjiajie, China, July 2012
Previous Work


Composability Verification & Validation in Component-based Model Development

Valid + Invalid Models
Composability Verification & Validation in Component-based Model Development

VALID (syntax ✓)

INVALID (syntax ✗)
Composability Verification & Validation in Component-based Model Development

Syntax ✓, Semantic ✗

Syntax ✗, Semantic ✓

Syntax ✗, Semantic ✗
Composability Verification & Validation in Component-based Model Development

Model Validity = Syntactic + Semantic

But what about emergent properties?
What is Emergence?

• “Manifested discernible regular formations and interesting irregular phenomena of actions or patterns” [4]
  – Static/dynamic; deterministic/stochastic

• Behavior of the whole that cannot be identified analytically from the parts, ..... [1]
  – Types of emergence: nominal, weak, strong
Classification of Causation-Based Emergence

Emergent system

Macro level

System states

Non-emergence

Emergence

No downward causation

Simple emergence

Deducible

Non-deducible

Simple deductible emergence

Simple observational emergence

Positive or negative DC

Complex emergence

Positive and negative DC

Complex DC

Weak emergence

Multiple emergence

Strong emergence

Weak stable emergence

Weak unstable emergence

Evolutionary emergence

Downward causation

Upward causation

Internal system
# Emergence Perspectives

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Main Problem</th>
<th>Key Ideas about Emergence</th>
<th>Examples</th>
<th>Issues</th>
</tr>
</thead>
</table>
| Philosophy      | How intelligence emerges from unintelligent matter?                          | - Limitations in our knowledge  
- Prediction                      | General                                                                                   | Cannot tell the causes of emergence                                                                 |
| Natural Sciences| **Biology**: How life emerges from inanimate matter?  
**Chemistry**: how a completely new substance emerges from other ones? | More than the sum, surprise, novelty  
Upward & downward causation   
Self-organization              | Bird flocking, ant colonies                                                               | Heavily depends on self-organization concept                                                                 |
| Social Sciences | How human behaviors emerges from the interactions between them                | Hierarchy  
Supervenience                                       | Crowd behavior                                                               | ??                                                                               |
| Computer Science| - Predict, detect, and validate emergent properties  
- Understand the causes of emergence                                          | Shifts in complexity  
- Disk drives performance  
- Priority inversion in O/S                                       | - Lack of emergent formalism  
- Emergence validation                                                         |
Philosophy’s Perspective

• Does not answer the following questions:
  – Where are emergent properties from?
  – What are the causes of emergence?

• Emergence cannot be understood without more advanced tools and knowledge
Emergence

• 2 levels & their interactions:
  – micro – individual components
  – macro – the system as a whole

• Key issues:
  – Identifying micro and macro levels
  – Defining micro-macro relationship to facilitate emergence identification
  – Identifying emergence as it happens, without previous knowledge
  – Determining if emergence is beneficial => emergence validation
  – ...

21 August 2014

PADS, Zhangjiajie, China, July 2012
Related Work

• Variable-based [22]
  – One or more variables describe emergent properties
  – The system run is recorded and relationships are deduced
  – Assumes prior knowledge about emergence

• Event-based [5]
  – Behavior as series of simple and complex events
  – Emergence defined as sequence of complex events
  – Assumes prior knowledge about emergence

• Grammar-based [16]
  – \( L_{\text{WHOLE}} \): properties of the system as a whole
  – \( L_{\text{PARTS}} \): properties obtained from the reunion of the parts
  – Emergence = \( L_{\text{WHOLE}} - L_{\text{PARTS}} \)
  – How to calculate \( L_{\text{PARTS}}, L_{\text{WHOLE}} \)?
## Related Work

<table>
<thead>
<tr>
<th>Formalism</th>
<th>Modeling Methods</th>
<th>Advantages</th>
<th>Issues</th>
</tr>
</thead>
</table>
| Variable-based     | Macroscopic models                      | Given a measure, it is usually not difficult to calculate it               | - Hard to choose a good measure  
| [22]               |                                         |                                                                            | - Does not gain much understanding of the causes of emergent properties                  |
| Event-based [5]    | Microscopic models                      |                                                                            | Need to define (formalize) event types and emergent behaviors beforehand                  |
| Grammar-based [16] | Cellular automata models, microscopic   | Better understanding of causes of emergence                               | Hard to define the sum operator for individual languages                                   |
|                    | models                                   |                                                                            |                                                                                            |
Proposed Approach

1. Identify emergent properties
   ✓ **Without prior knowledge** about emergence
   ✓ Representing components in terms of their *objectives*

2. Visualize and validate emergent behavior
   ✓ As it happens
Model Component Definition

- Component: abstracted as a *black-box* with *in* and/or *out* communication channel, and represented as a *meta-component*:
  - mandatory & specific attributes
  - behavior:
    - External – constraints on input and output (destination, origin, data type, range, etc.)
    - Internal – a finite timed state automaton

\[
[I_I]S_p[\Delta t] \xrightarrow{\text{Cond}_n} S_t[O_I][A_m]
\]

- Simulator: represented by *components* (base, model) linked using *connectors*
Objective-based Annotation

• Define components in terms of what it do, not how it is done

• Each component has an added objective $o(C_i)$
  – Defined using both textual and semantically-enriched variables
  – Type and range to signify that the objective is met

$$o(C_i) = \{(description, var) | var \in A_m\}$$

• Example: Birds in a Flock

  $o(b_i) = (description, var(o(b_i)))$

  $description = Fly \text{ northbound with an average speed of 20km/hour;}$

  $var(o(b_i)) = \{\text{direction, speed}\}$
Objective-based Approach

Component-based Model

(1) Individual Objectives & Isolated Components
(2) Reconstructability Analysis
(3) Calculated Composed Model State $S_r$

For each simulation step:

Calculate Distance $DS(S_r, S)$

Same State?

No → Emergent Behavior Set

(1) Simulation
(2) Record Simulation States
(3) Simulation State $S$
Reconstructability Analysis

• An approach for inducting modeling relationships and correlations between variables [2]

• Aim: to identify strongly related subsets of variables and to represent this knowledge in a simplified model which eliminates the connection between all other “almost" unrelated subsets of variables

• We used a variant called “reconstruction” – reverse process of constructing the whole from its parts
Calculating Composed Model State

- **Component state**, \( s(C_i)/o \), is defined as the reunion of all variables \( var \) that are defined in its objective \( o(C_i) \)

- **Composed model state**, \( S_r \), is the reunion of the states \( s(C_i)/o \) of components running in isolation

- At each simulation step
  - Add to \( S_r \) only states that are significantly different
**DS(S_r, S): Distance between States**

**Definition 1** (Semantic State Distance). Let \( s(p) = [\text{state}(C_1), \ldots, \text{state}(C_n)] \), \( s^*(q) = [\text{state}(C^*_1), \ldots, \text{state}(C^*_n)] \). The semantic state distance between vectors \( p \) and \( q \) is defined as

\[
DS(s(p), s^*(q)) = \frac{\sum_{i=1}^{n} |ds(\text{state}(C_i), \text{state}(C^*_i))|}{n}
\]

where \( ds(\text{state}(C_i), \text{state}(C^*_i)) = \frac{\sum_{a_i \in A(C_i), a^*_j \in A(C^*_j)} d(a_i, a^*_j)}{m} \), \( A(C_i) \) is the set of attributes for component \( C_i \), \( m = |A(C_i)| \) and \( d(a_i, a^*_j) \) is defined as

\[
d(a_i, a^*_j) = \begin{cases} 
0 & \text{if related}(a_i, a^*_j) \text{ and } \text{value}(a_i) = \text{value}(a^*_j) \\
0.5 & \text{if related}(a_i, a^*_j) \text{ and } \text{value}(a_i) \neq \text{value}(a^*_j) \\
1 & \text{if } \exists a^*_j \in A(C^*_i) \text{ s.t. related}(a_i, a^*_j) = \text{true}
\end{cases}
\]

where \( \text{related}(a_i, a_j) \) signifies that \( a_i \) and \( a_j \) are related in the COSMO ontology.
Example: Flock of Birds

• Each component abstracts a moving bird that changes its position
  – separation - avoid crowding
  – alignment - steer towards the average herding of local flockmates
  – cohesion - move towards the average position of local flockmates

• Objective

\[ o(b_i) = (\text{description}, \text{var}(o(b_i))) \]

\[ \text{description} = \text{Fly northbound with an average speed of 20km/hour;} \]
\[ \text{var}(o(b_i)) = \{\text{direction, speed}\} \]
Example: Flock of Birds
Snapshots of State Changes

### Composed Model (Reconstructability Analysis)

<table>
<thead>
<tr>
<th>Bird 1</th>
<th>Bird 2</th>
<th>Bird 3</th>
<th>Bird 4</th>
<th>Bird 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>120,N</td>
<td>114,N</td>
<td>135,N</td>
<td>111,N</td>
<td>98,N</td>
</tr>
<tr>
<td>155,N</td>
<td>193,N</td>
<td>146,N</td>
<td>180,N</td>
<td>150,N</td>
</tr>
<tr>
<td>360,N</td>
<td>394,N</td>
<td>387,N</td>
<td>380,N</td>
<td>350,N</td>
</tr>
<tr>
<td>539,N</td>
<td>571,N</td>
<td>560,N</td>
<td>558,N</td>
<td>538,N</td>
</tr>
<tr>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
</tr>
<tr>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
<td>-,N</td>
</tr>
</tbody>
</table>

### Real System (Simulation) (component interactions)

<table>
<thead>
<tr>
<th>Bird 1</th>
<th>Bird 2</th>
<th>Bird 3</th>
<th>Bird 4</th>
<th>Bird 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>53,S</td>
<td>209,N</td>
<td>128,N</td>
<td>67,S</td>
<td>43,N</td>
</tr>
<tr>
<td>45,S</td>
<td>174,NW</td>
<td>88,N</td>
<td>65,S</td>
<td>44,N</td>
</tr>
<tr>
<td>40,S</td>
<td>129,NW</td>
<td>59,N</td>
<td>60,SW</td>
<td>43,N</td>
</tr>
<tr>
<td>39,S</td>
<td>76,NW</td>
<td>42,N</td>
<td>51,SW</td>
<td>38,N</td>
</tr>
<tr>
<td>38,N</td>
<td>27,NW</td>
<td>48,N</td>
<td>34,NW</td>
<td>37,N</td>
</tr>
<tr>
<td>54,N</td>
<td>71,NW</td>
<td>69,NW</td>
<td>11,NW</td>
<td>57,N</td>
</tr>
</tbody>
</table>

<distance, direction>  
*where distance is relative to centre of drawing panel*

<-, N> – out of viewing range

21 August 2014  
PADS, Zhangjiajie, China, July 2012
Conclusions

• Two-step approach:
  – Identification of emergent properties without prior knowledge of emergence
  • Objective-based annotation of model components: what rather than how
  • Reconstructability analysis
  – Validation of emergent behavior

• First step in the identification of emergent properties
  – Limiting assumptions
  – Other examples: traffic junctions, social networks, ...
  – Visualization & Validation
Limitations

• Objective definition
  – Requires component creator’s knowledge
  – Dependent on how the component will be used

• Reconstructability analysis
  – Requires numerous iterations to converge to statistical significant differences

• Sensitivity analysis on threshold values
  – Until which point is emergent behavior normal and expected?

• ...
Thank you!

email: teoym@comp.nus.edu.sg, claudia.szabo@adelaide.edu.au

