Modeling Complex Systems: A Coalgebraic Perspective

Sun Meng

LMAM & Department of Informatics, School of Mathematical Sciences, Peking University, Beijing, China http://www.math.pku.edu.cn/teachers/sunm

November 29, 2016

Introduction	Components as Coalgebras	UML	Logic	Future Directions			
Roadmap							

- 1. Introduction.
- 2. Components as Coalgebras.
- 3. A Coalgebraic Perspective on UML.
- **4.** Coalgebras and Logic (almost \emptyset).
- 5. Future Opportunities.

- The development of science proceeds in a cycle of activities:
 - Analysis: Understand the problem;
 - Abstraction: creating a mathematical model by eliminating irrelevant details in order to identify what is essential;
 - Reasoning: reasoning within the model, getting a collection of general laws;
 - Specialization: the general laws are instantiated to the specific problem and a solution (= an implementation) is calculated, which leads to further understanding, and input for another round of activities.
- Is this process relevant to the development of software systems?

- The development of science proceeds in a cycle of activities:
 - Analysis: Understand the problem;
 - Abstraction: creating a mathematical model by eliminating irrelevant details in order to identify what is essential;
 - Reasoning: reasoning within the model, getting a collection of general laws;
 - Specialization: the general laws are instantiated to the specific problem and a solution (= an implementation) is calculated, which leads to further understanding, and input for another round of activities.
- Is this process relevant to the development of software systems?

- The development of science proceeds in a cycle of activities:
 - Analysis: Understand the problem;
 - Abstraction: creating a mathematical model by eliminating irrelevant details in order to identify what is essential;
 - Reasoning: reasoning within the model, getting a collection of general laws;
 - Specialization: the general laws are instantiated to the specific problem and a solution (= an implementation) is calculated, which leads to further understanding, and input for another round of activities.
- Is this process relevant to the development of software systems?

- The development of science proceeds in a cycle of activities:
 - Analysis: Understand the problem;
 - Abstraction: creating a mathematical model by eliminating irrelevant details in order to identify what is essential;
 - Reasoning: reasoning within the model, getting a collection of general laws;
 - Specialization: the general laws are instantiated to the specific problem and a solution (= an implementation) is calculated, which leads to further understanding, and input for another round of activities.
- Is this process relevant to the development of software systems?

- The development of science proceeds in a cycle of activities:
 - Analysis: Understand the problem;
 - Abstraction: creating a mathematical model by eliminating irrelevant details in order to identify what is essential;
 - Reasoning: reasoning within the model, getting a collection of general laws;
 - Specialization: the general laws are instantiated to the specific problem and a solution (= an implementation) is calculated, which leads to further understanding, and input for another round of activities.
- Is this process relevant to the development of software systems?



- The development of software systems proceeds in a cycle of activities:
 - Requirement Analysis: Understanding the domain problem;
 - System Modeling: Creating a model for the system by eliminating irrelevant details in order to identify essential properties;
 - Detailed Design: Providing detailed specification of the system;
 - Implementation: Final System;
 - Verification and Testing: To grarantee the correctness of the final implementation w.r.t. the specification.



- The development of software systems proceeds in a cycle of activities:
 - Requirement Analysis: Understanding the domain problem;
 - System Modeling: Creating a model for the system by eliminating irrelevant details in order to identify essential properties;
 - Detailed Design: Providing detailed specification of the system;
 - Implementation: Final System;
 - Verification and Testing: To grarantee the correctness of the final implementation w.r.t. the specification.



- The development of software systems proceeds in a cycle of activities:
 - Requirement Analysis: Understanding the domain problem;
 - System Modeling: Creating a model for the system by eliminating irrelevant details in order to identify essential properties;
 - Detailed Design: Providing detailed specification of the system;
 - Implementation: Final System;
 - Verification and Testing: To grarantee the correctness of the final implementation w.r.t. the specification.



- The development of software systems proceeds in a cycle of activities:
 - Requirement Analysis: Understanding the domain problem;
 - System Modeling: Creating a model for the system by eliminating irrelevant details in order to identify essential properties;
 - Detailed Design: Providing detailed specification of the system;
 - Implementation: Final System;
 - Verification and Testing: To grarantee the correctness of the final implementation w.r.t. the specification.



- The development of software systems proceeds in a cycle of activities:
 - Requirement Analysis: Understanding the domain problem;
 - System Modeling: Creating a model for the system by eliminating irrelevant details in order to identify essential properties;
 - Detailed Design: Providing detailed specification of the system;
 - Implementation: Final System;
 - Verification and Testing: To grarantee the correctness of the final implementation w.r.t. the specification.

Introduction

- Component based systems: The notion of components and the compositional design principle are well established in all other engineering disciplines, but until 1990s, were unsuccessful in the world of software systems.
- Component Software Technologies: CORBA (OMG), COM+ (MicroSoft), JavaBeans (Sun), ...
- What is a software component?
 - Software components are executable units of independent production, acquisition, and deployment that can be composed into a functioning system. (C. Szyperski, D. Gruntz and S.Murer 2003)
- The characteristic properties of a component are that it:
 - is a unit of independent deployment;
 - is a unit of third-party composition;
 - has no external observable state

Motivation: Question and Approach

- Component based systems: The notion of components and the compositional design principle are well established in all other engineering disciplines, but until 1990s, were unsuccessful in the world of software systems.
- Component Software Technologies: CORBA (OMG), COM+ (MicroSoft), JavaBeans (Sun), ...
- What is a software component?

- Software components are executable units of independent production, acquisition, and deployment that can be composed into a functioning system. (C. Szyperski, D. Gruntz and S.Murer 2003)
- The characteristic properties of a component are that it:
 - is a unit of independent deployment;
 - is a unit of third-party composition;
 - has no external observable state

Motivation: Question and Approach

- Component based systems: The notion of components and the compositional design principle are well established in all other engineering disciplines, but until 1990s, were unsuccessful in the world of software systems.
- Component Software Technologies: CORBA (OMG), COM+ (MicroSoft), JavaBeans (Sun), ...
- What is a software component?

- Software components are executable units of independent production, acquisition, and deployment that can be composed into a functioning system. (C. Szyperski, D. Gruntz and S.Murer 2003)
- The characteristic properties of a component are that it:
 - is a unit of independent deployment;
 - is a unit of third-party composition;
 - has no external observable state

Motivation: Question and Approach

- Component based systems: The notion of components and the compositional design principle are well established in all other engineering disciplines, but until 1990s, were unsuccessful in the world of software systems.
- Component Software Technologies: CORBA (OMG), COM+ (MicroSoft), JavaBeans (Sun), ...
- What is a software component?

- Software components are executable units of independent production, acquisition, and deployment that can be composed into a functioning system. (C. Szyperski, D. Gruntz and S.Murer 2003)
- The characteristic properties of a component are that it:
 - is a unit of independent deployment;
 - is a unit of third-party composition;
 - has no external observable state

Motivation: Question and Approach

- Component based systems: The notion of components and the compositional design principle are well established in all other engineering disciplines, but until 1990s, were unsuccessful in the world of software systems.
- Component Software Technologies: CORBA (OMG), COM+ (MicroSoft), JavaBeans (Sun), ...
- What is a software component?

- Software components are executable units of independent production, acquisition, and deployment that can be composed into a functioning system. (C. Szyperski, D. Gruntz and S.Murer 2003)
- The characteristic properties of a component are that it:
 - is a unit of independent deployment;
 - is a unit of third-party composition;
 - has no external observable state

Introduction

UML

- Could we apply the general approach to the development of component-based systems?
- **Key:** resort to the *algebra* vs. *coalgebra duality* as a mathematical explanation of the intuitive symmetry between data and behavioral structures.
 - Algebra: abstract description of data structures. The emphasis is on construction.
 - Coalgebra: abstract description of systems' behaviors. The emphasis is on observation.
- A mathematical model for components and their composition.
- Applying the model to component-based complex system modeling and design.

Introduction Components as Coalgebras UML Logic

- Could we apply the general approach to the development of component-based systems?
- **Key:** resort to the *algebra* vs. *coalgebra duality* as a mathematical explanation of the intuitive symmetry between data and behavioral structures.
 - Algebra: abstract description of data structures. The emphasis is on construction.
 - Coalgebra: abstract description of systems' behaviors. The emphasis is on observation.
- A mathematical model for components and their composition.
- Applying the model to component-based complex system modeling and design.

Components as Coalgebras UML Logic

Introduction

- Could we apply the general approach to the development of component-based systems?
- **Key:** resort to the *algebra* vs. *coalgebra duality* as a mathematical explanation of the intuitive symmetry between data and behavioral structures.
 - Algebra: abstract description of data structures. The emphasis is on construction.
 - Coalgebra: abstract description of systems' behaviors. The emphasis is on observation.
- A mathematical model for components and their composition.
- Applying the model to component-based complex system modeling and design.

Components as Coalgebras UML Logic Futu

Introduction

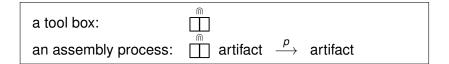
- Could we apply the general approach to the development of component-based systems?
- **Key:** resort to the *algebra* vs. *coalgebra duality* as a mathematical explanation of the intuitive symmetry between data and behavioral structures.
 - Algebra: abstract description of data structures. The emphasis is on construction.
 - Coalgebra: abstract description of systems' behaviors. The emphasis is on observation.
- A mathematical model for components and their composition.
- Applying the model to component-based complex system modeling and design.



• Algebra: abstract description of data structures.

```
[\mathsf{nil},\mathsf{cons}]:\mathbf{1}+\mathbf{A}\times\mathbf{A}^*\to\mathbf{A}^*
```

- The emphasis is on construction;
- In general:





Coalgebra: abstract description of systems' behaviors.

 $\langle \mathsf{head}, \mathsf{tail} \rangle : A^{\infty} \longrightarrow \mathbf{1} + A \times A^{\infty}$

- The emphasis is on observation;
- In general:

lens:	0~0
an observational structure:	universe $\stackrel{\rho}{\longrightarrow}$ \bigcirc \bigcirc universe



Functional Components

Question: What is the appropriate model for a component?

$f: I \longrightarrow O$

- The behavior of a function is captured by the output it produces, which is completely determined by the supplied input.
- Reality is not so simple!



• Question: What is the appropriate model for a component?

 $f:I\longrightarrow O$

- The behavior of a function is captured by the output it produces, which is completely determined by the supplied input.
- Reality is not so simple!

Introduction	Components as Coalgebras	UML	Logic	Future Directions				
Functional Components								

• Question: What is the appropriate model for a component?

 $f:I\longrightarrow O$

- The behavior of a function is captured by the output it produces, which is completely determined by the supplied input.
- Reality is not so simple!

• Is there any other possibilities?

• One may know how to produce output from input but not in all cases:

$$f: I \longrightarrow O+1$$

• One may be uncertain of the outcome of , in the sense that the evolution of the system being observed may be nondeterministic:

$$f: I \longrightarrow \mathscr{P}O$$

• One may recognize that there is some environmental (or context) information. (For example, it might be the case that the computation will modify the environment, thus influencing latter computation:

$$f: I \longrightarrow (O \times U)^U$$

- Is there any other possibilities?
- One may know how to produce output from input but not in all cases:

$$f: I \longrightarrow O + 1$$

• One may be uncertain of the outcome of , in the sense that the evolution of the system being observed may be nondeterministic:

$$f: I \longrightarrow \mathscr{P}O$$

• One may recognize that there is some environmental (or context) information. (For example, it might be the case that the computation will modify the environment, thus influencing latter computation:

$$f: I \longrightarrow (O \times U)^U$$

- Is there any other possibilities?
- One may know how to produce output from input but not in all cases:

$$f: I \longrightarrow O + 1$$

 One may be uncertain of the outcome of , in the sense that the evolution of the system being observed may be nondeterministic:

$$f: I \longrightarrow \mathscr{P}O$$

• One may recognize that there is some environmental (or context) information. (For example, it might be the case that the computation will modify the environment, thus influencing latter computation:

$$f: I \longrightarrow (O \times U)^U$$

- Is there any other possibilities?
- One may know how to produce output from input but not in all cases:

$$f: I \longrightarrow O + 1$$

 One may be uncertain of the outcome of , in the sense that the evolution of the system being observed may be nondeterministic:

$$f: I \longrightarrow \mathscr{P}O$$

 One may recognize that there is some environmental (or context) information. (For example, it might be the case that the computation will modify the environment, thus influencing latter computation:

$$f: I \longrightarrow (O \times U)^U$$

- A function computed within a context is often referred to as state-based, in the sense the word 'state' has in automata theory - the internal memory of the automata which both constraints and is constrained by the execution of actions.
- The 'nature' of *f* : *I* → (*O*×*U*)^{*U*} as a 'state-based function' is made more explicit by rewriting it as

$$f: U \longrightarrow (O \times U)^{T}$$

• One's focus becomes the 'universe' or, more pragmatically, the state space. Input and output parameters may or may not be relevant, depending on the particular kind of observation one may want to perform.



- A function computed within a context is often referred to as state-based, in the sense the word 'state' has in automata theory - the internal memory of the automata which both constraints and is constrained by the execution of actions.
- The 'nature' of *f* : *I* → (*O*×*U*)^{*U*} as a 'state-based function' is made more explicit by rewriting it as

$$f: U \longrightarrow (O \times U)^{\prime}$$

• One's focus becomes the 'universe' or, more pragmatically, the state space. Input and output parameters may or may not be relevant, depending on the particular kind of observation one may want to perform.

- A function computed within a context is often referred to as state-based, in the sense the word 'state' has in automata theory - the internal memory of the automata which both constraints and is constrained by the execution of actions.
- The 'nature' of *f* : *I* → (*O*×*U*)^{*U*} as a 'state-based function' is made more explicit by rewriting it as

$$f: U \longrightarrow (O \times U)^{I}$$

 One's focus becomes the 'universe' or, more pragmatically, the state space. Input and output parameters may or may not be relevant, depending on the particular kind of observation one may want to perform.

- Informal understanding of components shows that it has:
 - internal state space that persists in time;
 - the possibility of interaction with other components during the overall computation;
 - observable through well-defined interfaces to ensure flow of data.
- Such components can be found "everywhere", from sophisticated plant control systems, to formal automata or domestic appliances.
- To investigate components one should equip himself with an appropriate 'lens' _____ which necessarily entails a particular shape for observation.

- Informal understanding of components shows that it has:
 - internal state space that persists in time;
 - the possibility of interaction with other components during the overall computation;
 - observable through well-defined interfaces to ensure flow of data.
- Such components can be found "everywhere", from sophisticated plant control systems, to formal automata or domestic appliances.
- To investigate components one should equip himself with an appropriate 'lens' _____ which necessarily entails a particular shape for observation.

- Informal understanding of components shows that it has:
 - internal state space that persists in time;
 - the possibility of interaction with other components during the overall computation;
 - observable through well-defined interfaces to ensure flow of data.
- Such components can be found "everywhere", from sophisticated plant control systems, to formal automata or domestic appliances.

• A lens _____ for components is an operation mapping a set (of states) *U* to a set _____ *U* containing the possible effects of an observable transition.

0~0	$\bigcirc \frown \bigcirc U$	Component Behavior		
1	1	stop		
A	A	outputs $a \in A$ once		
ld	U	running forever		
$A \times -$	A imes U	stream over A		
$A \times -+1$	$A \times U + 1$	finite or infinite list over A		

• Or provides an appropriate notion of interface for components.

• A lens _____ for components is an operation mapping a set (of states) *U* to a set _____ *U* containing the possible effects of an observable transition.

0~0	$\bigcirc \frown \bigcirc U$	Component Behavior		
1	1	stop		
A	A	outputs $a \in A$ once		
ld	U	running forever		
$A \times -$	A imes U	stream over A		
$A \times -+1$	$A \times U + 1$	finite or infinite list over A		

• Or provides an appropriate notion of interface for components.

 The behavior of the sort of computational structures is determined not only by an external input stimuli, but also by some internal 'memory' to which there is, in general, no direct access. Such systems can always be represented by functions typed as

 $p: U \longrightarrow \bigcirc \bigcirc \bigcirc U$

- Interpretation:
 - U: the set of states;
 - *p*: the component's dynamics, describing the observable effects of an elementary step (transition) in the evolution of the component;
 - O____: the shape for the observation structure (*interface*);
 - Or U: the set of all possible outcomes of taking one step transition.

- Mealy machines (on and) were introduced as components observed through the interface (lens) ○ ○ = (O × −)^l.
- Different interfaces give rise to different classes of components.

- Mealy machines (on and) were introduced as components observed through the interface (lens) $\bigcirc \bigcirc \bigcirc = (O \times -)^{I}$.
- can be regarded as the type of a mapping which decomposes the 'observable universe' U into an 'observation context' $(O \times U)^{I}$.

- Mealy machines (on and) were introduced as components observed through the interface (lens) $\bigcirc \bigcirc \bigcirc = (O \times -)^{I}$.
- can be regarded as the type of a mapping which decomposes the 'observable universe' U into an 'observation context' $(O \times U)^{I}$.
- Different interfaces give rise to different classes of compo-• nents.

- Interface shapes for observations:
 - 'opaque': ∪ = 1
 - black and white: $\bigcirc \bigcirc U = 2$
 - colouring: $\bigcirc \bigcirc U = O$
 - multi-attribute: $\bigcirc \frown \bigcirc U = \prod_{k \in K} O_k$
- Interface shapes for actions:
 - partiality: $\bigcirc \frown \bigcirc U = U + 1$
 - visible attributes (outputs): $\bigcirc \frown \bigcirc U = O \times U$
 - external stimulus triggered evolution (inputs): $\bigcirc \bigcirc U = U^{I}$
 - non-determinism: $\bigcirc \frown \bigcirc U = \mathscr{P}U$
 - probability: $\bigcirc \bigcirc U = \mathscr{D}U$, where for a set $X, \mathscr{D}X = \{\xi : X \to [0,1] \mid \sum_{x \in X} \xi(x) \le 1\}$, and ξ is called a probability sub-distribution over X.
- The elementary 'lens' can be glued with set-theoretic constructions.

UML

- Interface shapes for observations:
 - 'opaque': ∪ = 1
 - black and white: $\bigcirc \bigcirc U = 2$
 - colouring: $\bigcirc \bigcirc U = O$
 - multi-attribute: $\bigcirc \frown \bigcirc U = \prod_{k \in K} O_k$
- Interface shapes for actions:
 - partiality: $\bigcirc \bigcirc U = U + \mathbf{1}$
 - visible attributes (outputs): $\bigcirc \bigcirc U = O \times U$
 - external stimulus triggered evolution (inputs): $\bigcirc \bigcirc U = U^{I}$
 - non-determinism: $\bigcirc \bigcirc U = \mathscr{P}U$
 - probability: $\bigcirc \bigcirc U = \mathscr{D}U$, where for a set $X, \mathscr{D}X = \{\xi : X \to [0,1] \mid \sum_{x \in X} \xi(x) \le 1\}$, and ξ is called a probability sub-distribution over X.
- The elementary 'lens' can be glued with set-theoretic constructions.

UML

- Interface shapes for observations:
 - 'opaque': ∪ = 1
 - black and white: $\bigcirc \bigcirc U = 2$
 - colouring: $\bigcirc \bigcirc U = O$
 - multi-attribute: $\bigcirc \frown \bigcirc U = \prod_{k \in K} O_k$
- Interface shapes for actions:
 - partiality: $\bigcirc \bigcirc U = U + \mathbf{1}$
 - visible attributes (outputs): $\bigcirc \bigcirc U = O \times U$
 - external stimulus triggered evolution (inputs): $\bigcirc \bigcirc U = U^{I}$
 - non-determinism: $\bigcirc \bigcirc U = \mathscr{P}U$
 - probability: $\bigcirc \bigcirc U = \mathscr{D}U$, where for a set $X, \mathscr{D}X = \{\xi : X \to [0,1] \mid \sum_{x \in X} \xi(x) \le 1\}$, and ξ is called a probability sub-distribution over X.
- The elementary 'lens' can be glued with set-theoretic constructions.



- For a universe U and observation structure p : U → ○○○U, the pair (U, p) is called a ○○○-coalgebra.
- A morphism between coalgebras is a function between their carriers which preserves the dynamics, i.e.

$$U \xrightarrow{p} \bigcirc \bigcup U$$

$$h \downarrow \qquad \qquad \downarrow \bigcirc \frown \bigcirc h$$

$$V \xrightarrow{q} \bigcirc \bigcup V$$



- O should be applicable not only to *sets*, but also to *functions*.
- The idea of an uniform transformation of both sets and functions is captured by the notion of a **functor**.
- A functor is a function over our working universe which preserves *identities* and *composition*, i.e., the graph and monoidal structure:
 - For each function $f : A \rightarrow B$, $Tf : TA \rightarrow TB$;
 - $\mathsf{T}id_X = id_{\mathsf{T}X};$
 - $T(f \circ g) = Tf \circ Tg$.
- The functor should capture both a signature of actions and observers, as well as a particular behavior model.



- O should be applicable not only to *sets*, but also to *functions*.
- The idea of an uniform transformation of both sets and functions is captured by the notion of a **functor**.
- A functor is a function over our working universe which preserves *identities* and *composition*, i.e., the graph and monoidal structure:
 - For each function $f : A \rightarrow B$, $Tf : TA \rightarrow TB$;
 - Tid_X = id_{TX};
 - $T(f \circ g) = Tf \circ Tg$.
- The functor should capture both a signature of actions and observers, as well as a particular behavior model.



- O should be applicable not only to *sets*, but also to *functions*.
- The idea of an uniform transformation of both sets and functions is captured by the notion of a **functor**.
- A functor is a function over our working universe which preserves *identities* and *composition*, i.e., the graph and monoidal structure:
 - For each function $f : A \rightarrow B$, $Tf : TA \rightarrow TB$;
 - $Tid_X = id_{TX};$
 - $T(f \circ g) = Tf \circ Tg$.
- The functor should capture both a signature of actions and observers, as well as a particular behavior model.



- Or should be applicable not only to *sets*, but also to *functions*.
- The idea of an uniform transformation of both sets and functions is captured by the notion of a **functor**.
- A functor is a function over our working universe which preserves *identities* and *composition*, i.e., the graph and monoidal structure:
 - For each function $f : A \rightarrow B$, $Tf : TA \rightarrow TB$;
 - $Tid_X = id_{TX};$
 - $T(f \circ g) = Tf \circ Tg$.
- The functor should capture both a signature of actions and observers, as well as a particular behavior model.

- These aspects are orthogonal and should be dealt separately.
- Therefore we consider the interface functor

$$\mathsf{T}^{\mathsf{B}}_{l,O} = \mathsf{B}(\mathsf{Id} \times O)^{l}$$

where *I* and *O* are sets acting as component input and output interfaces.

• For a component *p* with such an interface, the transition structure of the corresponding coalgebra can be represented as

$$\overline{\alpha_p}: U_p \longrightarrow \mathsf{B}(U_p \times O)^T$$

where $\alpha_p : U_p \times I \rightarrow B(U_p \times O)$ represents the state transitions.

A component *p* for interface T^B_{l,O} can be represented as a seeded coalgebra *p* = ⟨U_p, α_p : U_p → T^B_{l,O}U_p, u₀⟩.

- These aspects are orthogonal and should be dealt separately.
- Therefore we consider the interface functor

$$\mathsf{T}^{\mathsf{B}}_{I,O} = \mathsf{B}(\mathsf{Id} \times O)^{I}$$

where *I* and *O* are sets acting as component input and output interfaces.

• For a component *p* with such an interface, the transition structure of the corresponding coalgebra can be represented as

$$\overline{\alpha_p}: U_p \longrightarrow \mathsf{B}(U_p \times O)^T$$

where $\alpha_p : U_p \times I \rightarrow B(U_p \times O)$ represents the state transitions.

A component *p* for interface T^B_{l,O} can be represented as a seeded coalgebra *p* = ⟨U_p, α_p : U_p → T^B_{l,O}U_p, u₀⟩.

- These aspects are orthogonal and should be dealt separately.
- Therefore we consider the interface functor

$$\mathsf{T}^{\mathsf{B}}_{I,O} = \mathsf{B}(\mathsf{Id} \times O)^{I}$$

where *I* and *O* are sets acting as component input and output interfaces.

• For a component *p* with such an interface, the transition structure of the corresponding coalgebra can be represented as

$$\overline{\alpha_{p}}: U_{p} \longrightarrow \mathsf{B}(U_{p} \times O)^{T}$$

where $\alpha_p : U_p \times I \rightarrow B(U_p \times O)$ represents the state transitions.

A component *p* for interface T^B_{l,O} can be represented as a seeded coalgebra *p* = ⟨U_p, α_p : U_p → T^B_{l,O}U_p, u₀⟩.

- These aspects are orthogonal and should be dealt separately.
- Therefore we consider the interface functor

$$\mathsf{T}^{\mathsf{B}}_{I,O} = \mathsf{B}(\mathsf{Id} \times O)^{I}$$

where *I* and *O* are sets acting as component input and output interfaces.

• For a component *p* with such an interface, the transition structure of the corresponding coalgebra can be represented as

$$\overline{\alpha_{p}}: U_{p} \longrightarrow \mathsf{B}(U_{p} \times O)^{T}$$

where $\alpha_p : U_p \times I \rightarrow B(U_p \times O)$ represents the state transitions.

A component *ρ* for interface T^B_{l,O} can be represented as a seeded coalgebra *ρ* = ⟨*U_ρ*, *α_ρ* : *U_ρ* → T^B_{l,O}*U_ρ*, *u*₀⟩.

- Successive observations of a component *p* reveal its allowed behavioral patterns.
- For each state value *u* ∈ *U_p*, the behavior of *p* at *u* (more precisely, from *u* onwards) organize itself into a tree-like structure, because it depends on the sequences of input items processed.
- Such trees, whose arcs are labelled with *I* values and nodes with *O* values, can be represented by functions from non empty sequence of *I* to B-structures of output items.
- In other words, the space of behaviors of a component with input *I* and output *O* is the set (B*O*)^{*I*⁺}, which is in fact the carrier ν_T of the final T-coalgebra (ν_T, ω_T : ν_T → Tν_T).

- Successive observations of a component *p* reveal its allowed behavioral patterns.
- For each state value *u* ∈ *U_p*, the behavior of *p* at *u* (more precisely, from *u* onwards) organize itself into a tree-like structure, because it depends on the sequences of input items processed.
- Such trees, whose arcs are labelled with *I* values and nodes with *O* values, can be represented by functions from non empty sequence of *I* to B-structures of output items.
- In other words, the space of behaviors of a component with input *I* and output *O* is the set (B*O*)^{*I*⁺}, which is in fact the carrier ν_T of the final T-coalgebra (ν_T, ω_T : ν_T → Tν_T).

- Successive observations of a component *p* reveal its allowed behavioral patterns.
- For each state value *u* ∈ *U_p*, the behavior of *p* at *u* (more precisely, from *u* onwards) organize itself into a tree-like structure, because it depends on the sequences of input items processed.
- Such trees, whose arcs are labelled with *I* values and nodes with *O* values, can be represented by functions from non empty sequence of *I* to B-structures of output items.
- In other words, the space of behaviors of a component with input *I* and output *O* is the set (B*O*)^{*I*+}, which is in fact the carrier ν_T of the final T-coalgebra (ν_T, ω_T : ν_T → Tν_T).

- Successive observations of a component *p* reveal its allowed behavioral patterns.
- For each state value *u* ∈ *U_p*, the behavior of *p* at *u* (more precisely, from *u* onwards) organize itself into a tree-like structure, because it depends on the sequences of input items processed.
- Such trees, whose arcs are labelled with *I* values and nodes with *O* values, can be represented by functions from non empty sequence of *I* to B-structures of output items.
- In other words, the space of behaviors of a component with input *I* and output *O* is the set (B*O*)^{*I*⁺}, which is in fact the carrier ν_T of the final T-coalgebra (ν_T, ω_T : ν_T → Tν_T).



• By finality, from any other T-coalgebra *p*, there is a unique morphism [(*p*)] making the following diagram to commute:

$$\begin{array}{c} \nu_{\mathsf{T}} \xrightarrow{\omega_{\mathsf{T}}} \mathsf{B}(\nu_{\mathsf{T}} \times O)' \\ \llbracket(\rho)\rrbracket^{\dagger} & \uparrow \mathsf{B}(\llbracket(\rho)\rrbracket \times O)' \\ U_{\rho} \xrightarrow{\overline{\alpha}_{\rho}} \mathsf{B}(U_{\rho} \times O)' \end{array}$$

Applying morphism [(*p*)] to a state value *u* ∈ *U_p* gives the observable behavior of a sequence of *p* transitions starting at *u*.

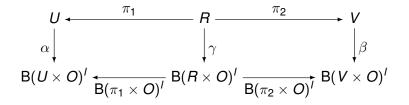
UML

Bisimulation

Bisimulation for two T-coalgebras (U, α) and (V, β) is a relation $R \subseteq U \times V$ such that there is a T-coalgebra (R, γ) satisfying

 $T(\pi_1) \circ \gamma = \alpha \circ \pi_1$ $T(\pi_2) \circ \gamma = \beta \circ \pi_2$

The following diagram is the corresponding instantiation for the functor T underlying our model of components.

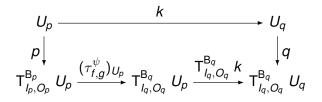


Introduction	Components as Coalgebras	UML	Logic	Future Directions
	Bisim	ulation		

- Provides a 'relational' view of coalgebra morphisms, as the graph of a T-coalgebra morphism is a T-bisimulation.
- Has a large theory (e.g. closed under converse and composition).
- Entails a *local* proof theory for observational equivalence: the *coinductive* proof principle being widely used in the coalgebra literature is the explicit construction of a bisimulation containing the pair of states one want to prove equivalent.
- Parametric on system's interface T.

- For every functor T, the T-coalgebras together with the homomorphisms between them form a category.
- A new category **Co** is defined as a "total category" which contains the seeded coalgebras for different functors.
- For two components *p* = ⟨*U_p*, *α_p* : *U_p* → T^{B_p}<sub>*l_p*,*O_p*, *U_p*, *u_p*⟩ and *q* = ⟨*U_q*, *α_q* : *U_q* → T^{B_q}_{*l_q*,*O_q}<i>U_q*, *u_q*⟩ in **Co**, the morphism between them is ⟨*k*, *τ^ψ_{f,g}*⟩ where *f* : *I_p* → *I_q*, *g* : *O_p* → *O_q*, *ψ* : B_p ⇒ B_q, *τ^ψ_{f,g}* = *ψ*(id × *g*)^{*f*} : T^{B_p}<sub>*l_p*,*O_p* ⇒ T^{B_q}<sub>*l_q*,*O_q* is a natural transformation, and *k* : *U_p* → *U_q* is the coalgebra morphism under natural transformation *τ^ψ_{f,g}*, such that *k u_p* = *u_q* and the following diagram commutes:
 </sub></sub></sub></sub>

Introduction	Components as Coalgebras	UML	Logic	Future Directions



Bisimulation of Components

Definition

Two T-components $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow T \ U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow T \ U_q, u_q \rangle$ are bisimilar iff there is a T-bisimulation containing the pair $\langle u_p, u_q \rangle$.

Definition

Let $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow \mathsf{T}_{l_p, \mathcal{O}_p}^{\mathsf{B}_p} U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow \mathsf{T}_{l_q, \mathcal{O}_q}^{\mathsf{B}_q} U_q, u_q \rangle$ be two components, and $\tau_{f,g}^{\psi} = \psi(\mathsf{id} \times g)^f$ is a natural transformation, where $f : I_p \longrightarrow I_q$, $g : \mathcal{O}_p \longrightarrow \mathcal{O}_q, \psi : \mathsf{B}_p \Rightarrow \mathsf{B}_q$, if $p' = \langle U_p, (\tau_{f,g}^{\psi})_{U_p} \cdot \overline{\alpha}_p, u_p \rangle$ and qare bisimilar, then p and q are bisimilar under the natural transformation $\tau_{f,g}^{\psi}$, denoted by $p \approx_{\tau_{f,q}^{\psi}} q$.

Bisimulation of Components

Definition

Two T-components $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow T \ U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow T \ U_q, u_q \rangle$ are bisimilar iff there is a T-bisimulation containing the pair $\langle u_p, u_q \rangle$.

Definition

Let $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow \mathsf{T}_{l_p, O_p}^{\mathsf{B}_p} U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow \mathsf{T}_{l_q, O_q}^{\mathsf{B}_q} U_q, u_q \rangle$ be two components, and $\tau_{f,g}^{\psi} = \psi(\mathsf{id} \times g)^f$ is a natural transformation, where $f : I_p \longrightarrow I_q$, $g : O_p \longrightarrow O_q, \psi : \mathsf{B}_p \Rightarrow \mathsf{B}_q$, if $p' = \langle U_p, (\tau_{f,g}^{\psi})_{U_p} \cdot \overline{\alpha}_p, u_p \rangle$ and qare bisimilar, then p and q are bisimilar under the natural transformation $\tau_{f,g}^{\psi}$, denoted by $p \approx_{\tau_{f,q}^{\psi}} q$.

Definition

For two components $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow \mathsf{T}_{I,K}^{\mathsf{B}_p} U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow \mathsf{T}_{K,O}^{\mathsf{B}_q} U_q, u_q \rangle$, the sequential composition of *p* and *q* is defined as

$${m
ho}$$
 ; ${m q}=\langle {m U},\overlinelpha_{{m
ho};{m q}}:{m U} \longrightarrow {\sf T}^{\sf B}_{I,{m O}} {m U},{m u}_{{m
ho};{m q}}
angle$

where $\textit{U}=\textit{U}_{p} imes \textit{U}_{q}, \textit{B}=\textit{B}_{p}\textit{B}_{q}, \textit{u}_{p;q}=\langle\textit{u}_{p},\textit{u}_{q}
angle$ and

$$\begin{array}{l} \alpha_{p;q} = U_p \times U_q \times I \xrightarrow{\text{xr}} U_p \times I \times U_q \xrightarrow{\alpha_p \times \text{id}} \mathsf{B}_p(U_p \times K) \times U_q \\ \xrightarrow{\tau_r} \mathsf{B}_p(U_p \times K \times U_q) \xrightarrow{\mathsf{B}_p(\mathfrak{a} \cdot \mathsf{xr})} \mathsf{B}_p(U_p \times (U_q \times K)) \\ \xrightarrow{\mathsf{B}_p(\text{id} \times \alpha_q)} \mathsf{B}_p(U_p \times \mathsf{B}_q(U_q \times O)) \xrightarrow{\mathsf{B}_p \tau_l} \mathsf{B}_p \mathsf{B}_q(U_p \times (U_q \times O)) \\ \xrightarrow{\mathsf{B}_p \mathsf{B}_q \mathfrak{a}^\circ} \mathsf{B}_p \mathsf{B}_q(U_p \times U_q \times O) = \mathsf{B}(U \times O) \end{array}$$

Definition

For two components $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow \mathsf{T}_{l_p, O_p}^{\mathsf{B}_p} U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow \mathsf{T}_{l_q, O_q}^{\mathsf{B}_q} U_q, u_q \rangle$, their external choice is defined as $p \boxplus q = \langle U, \overline{\alpha}_{p \boxplus q} : U \longrightarrow \mathsf{T}_{l, O}^{\mathsf{B}_{p \boxplus q}} U, u_{p \boxplus q} \rangle$, where $U = U_p \times U_q$, $I = I_p + I_q$, $O = O_p + O_q$, $\mathsf{B}_{p \boxplus q} = \mathsf{B}_p + \mathsf{B}_q$, $u_{p \boxplus q} = \langle u_p, u_q \rangle$, and

$$\begin{array}{l} \alpha_{p \boxplus q} = U \times (I_{p} + I_{q}) \xrightarrow{\mathrm{dr}} (U \times I_{p}) + (U \times I_{q}) \\ \xrightarrow{\mathrm{xr} + \mathbf{a}} (U_{p} \times I_{p}) \times U_{q} + U_{p} \times (U_{q} \times I_{q}) \\ \xrightarrow{\alpha_{p} \times \mathrm{id} + \mathrm{id} \times \alpha_{q}} & \mathsf{B}_{p} (U_{p} \times O_{p}) \times U_{q} + U_{p} \times \mathsf{B}_{q} (U_{q} \times O_{q}) \\ \xrightarrow{\tau_{r} + \tau_{l}} & \mathsf{B}_{p} ((U_{p} \times O_{p}) \times U_{q}) + \mathsf{B}_{q} (U_{p} \times (U_{q} \times O_{q})) \\ \xrightarrow{\mathsf{B}_{p} \times \mathrm{xr} + \mathsf{B}_{q} a^{\circ}} & \mathsf{B}_{p} (U \times O_{p}) + \mathsf{B}_{q} (U \times O_{q}) \\ \xrightarrow{[\mathsf{B}_{p} \iota_{1}, \mathsf{B}_{q} \iota_{2}]} & \mathsf{B}_{p \boxplus q} (U \times O) \end{array}$$

Definition

For two components $p = \langle U_p, \overline{\alpha}_p : U_p \longrightarrow \mathsf{T}_{l_p, O_p}^{\mathsf{B}_p} \ U_p, u_p \rangle$ and $q = \langle U_q, \overline{\alpha}_q : U_q \longrightarrow \mathsf{T}_{l_q, O_q}^{\mathsf{B}_q} \ U_q, u_q \rangle$, their parallel composition $p \boxtimes q = \langle U, \overline{\alpha}_{p\boxtimes q} : U \longrightarrow \mathsf{T}_{l, O}^{\mathsf{B}_{p\boxtimes q}} U, u_{p\boxtimes q} \rangle$ where $U = U_p \times U_q$, $I = I_p \times I_q, \ O = O_p \times O_q, \ B_{p\boxtimes q} = \mathsf{B}_p \ \mathsf{B}_q, \ u_{p\boxtimes q} = \langle u_p, u_q \rangle$ and $\alpha_{p\boxtimes q} = U_p \times U_q \times (I_p \times I_q) \xrightarrow{\mathsf{m}} (U_p \times I_p) \times (U_q \times I_q)$ $\xrightarrow{\alpha_p \times \alpha_q} \mathsf{B}_p \ (U_p \times O_p) \times \mathsf{B}_q \ (U_q \times O_q)$

$$\xrightarrow{\prime \prime} \mathsf{B}_{\rho} \left(U_{\rho} \times O_{\rho} \times \mathsf{B}_{q} \left(U_{q} \times O_{q} \right) \right)$$

$$\xrightarrow{\mathsf{B}_p \mid I_1} \mathsf{B}_p \mathsf{B}_q (U_p \times O_p \times (U_q \times O_q))$$

 $\xrightarrow{\mathsf{B}_{\rho} \mathsf{B}_{q} \mathsf{m}} \mathsf{B}_{\rho} \mathsf{B}_{q} \left(\mathit{U}_{\rho} \times \mathit{U}_{q} \times (\mathit{O}_{\rho} \times \mathit{O}_{q}) \right) \; = \; \mathsf{B}_{\rho \boxtimes q} \left(\mathit{U} \times \mathit{O} \right)$

Introduction	Components as Coalgebras	UML	Logic	Future Directions		
The Calculus of Components						

For a component $p = \langle U, \overline{\alpha}_p : U \longrightarrow \mathsf{T}^{\mathsf{B}}_{I,O} U, u_0 \rangle$ and functions $f : I' \longrightarrow I, g : O \longrightarrow O'$, the wrapping

$${\mathcal p}[f,g] = \langle U, \overline{lpha}_{{\mathcal p}[f,g]} : U \longrightarrow {\mathsf T}^{\mathsf B}_{I',{\mathcal O}'} \; U, u_0
angle$$

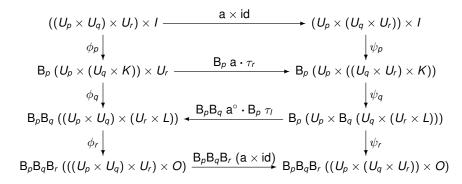
where

$$\begin{array}{l} \alpha_{p[f,g]} = U \times I' \xrightarrow{\mathsf{id} \times f} U \times I \xrightarrow{\alpha_p} \\ \mathsf{B}(U \times O) \xrightarrow{\mathsf{B}(\mathsf{id} \times g)} \mathsf{B}(U \times O') \end{array}$$

ntroduction	Components as Coalgebras	UML	Logic	Future Directions

$$\begin{array}{ll} (p;q);r \approx p;(q;r) & (1) \\ (p\boxtimes q)\boxtimes r \approx (p\boxtimes (q\boxtimes r))[a,a^{\circ}] & (2) \\ q\boxtimes p \approx_{\tau_{s,s}^{\gamma}} p\boxtimes q & (3) \\ \text{idle} \boxtimes p \approx_{\tau_{id,id}} p[r,r^{\circ}] & (4) \\ \text{nil} \boxtimes p \approx_{\tau_{id,id}} nil[zl,zl^{\circ}] & (5) \\ (p\boxplus q)\boxplus r \approx (p\boxplus (q\boxplus r))[a_{+},a^{\circ}] & (6) \\ p\boxplus q \approx_{\tau_{s+s}^{s+}} q\boxplus p & (7) \\ \text{nil} \boxplus p \approx_{\tau_{id,\pi_{2}}} p[r_{+},r^{\circ}] & (8) \\ p\boxplus nil \approx_{\tau_{id,\pi_{1}}} p[l_{+},l^{\circ}] & (9) \end{array}$$

Introduction	Components as Coalgebras	UML	Logic	Future Directions	
The Calculus of Components					





- Refinement: A *transformation* of an "abstract" into a more "concrete" design, entailing a notion of *substitution*.
- Data refinement, being traced back to Hoare's work, retrieve function from the concrete into the abstract model is defined.
- Object-orientation, substitution is expressed in terms of behavior typing.
- Process algebra, reduction of nondeterminism.
- Semantic characterization of refinement for state-based components.



- Refinement: A *transformation* of an "abstract" into a more "concrete" design, entailing a notion of *substitution*.
- Data refinement, being traced back to Hoare's work, retrieve function from the concrete into the abstract model is defined.
- Object-orientation, substitution is expressed in terms of behavior typing.
- Process algebra, reduction of nondeterminism.
- Semantic characterization of refinement for state-based components.



- Refinement: A *transformation* of an "abstract" into a more "concrete" design, entailing a notion of *substitution*.
- Data refinement, being traced back to Hoare's work, retrieve function from the concrete into the abstract model is defined.
- Object-orientation, substitution is expressed in terms of behavior typing.
- Process algebra, reduction of nondeterminism.
- Semantic characterization of refinement for state-based components.



- Refinement: A *transformation* of an "abstract" into a more "concrete" design, entailing a notion of *substitution*.
- Data refinement, being traced back to Hoare's work, retrieve function from the concrete into the abstract model is defined.
- Object-orientation, substitution is expressed in terms of behavior typing.
- Process algebra, reduction of nondeterminism.
- Semantic characterization of refinement for state-based components.



- Refinement: A *transformation* of an "abstract" into a more "concrete" design, entailing a notion of *substitution*.
- Data refinement, being traced back to Hoare's work, retrieve function from the concrete into the abstract model is defined.
- Object-orientation, substitution is expressed in terms of behavior typing.
- Process algebra, reduction of nondeterminism.
- Semantic characterization of refinement for state-based components.

Refinement

- Based on the coalgebraic framework, three kinds of refinement relations can be defined for state-based systems:
- *Behavioral Refinement*: typically relates systems of the same interface, where the refinement is based on a simulation preorder between the two systems.
- *Interface Refinement*: relates systems of different interfaces, and the question is whether a system can be transformed, by suitable wiring, to replace another system with a different interface.
- *Architectural Refinement*: being used for decomposing a system with a specified behavior into a distributed system architecture, i.e., a family of systems (components) combined in parallel.

Introduction	Components as Coalgebras	UML	Logic	Future Directions	
Refinement					

- Based on the coalgebraic framework, three kinds of refinement relations can be defined for state-based systems:
- *Behavioral Refinement*: typically relates systems of the same interface, where the refinement is based on a simulation preorder between the two systems.
- *Interface Refinement*: relates systems of different interfaces, and the question is whether a system can be transformed, by suitable wiring, to replace another system with a different interface.
- Architectural Refinement: being used for decomposing a system with a specified behavior into a distributed system architecture, i.e., a family of systems (components) combined in parallel.

Introduction	Components as Coalgebras	UML	Logic	Future Directions	
Refinement					

- Based on the coalgebraic framework, three kinds of refinement relations can be defined for state-based systems:
- *Behavioral Refinement*: typically relates systems of the same interface, where the refinement is based on a simulation preorder between the two systems.
- Interface Refinement: relates systems of different interfaces, and the question is whether a system can be transformed, by suitable wiring, to replace another system with a different interface.
- Architectural Refinement: being used for decomposing a system with a specified behavior into a distributed system architecture, i.e., a family of systems (components) combined in parallel.

Introduction	Components as Coalgebras	UML	Logic	Future Directions
	Refir	ement		

- Based on the coalgebraic framework, three kinds of refinement relations can be defined for state-based systems:
- *Behavioral Refinement*: typically relates systems of the same interface, where the refinement is based on a simulation preorder between the two systems.
- Interface Refinement: relates systems of different interfaces, and the question is whether a system can be transformed, by suitable wiring, to replace another system with a different interface.
- Architectural Refinement: being used for decomposing a system with a specified behavior into a distributed system architecture, i.e., a family of systems (components) combined in parallel.

- Behavior refinement affects the internal dynamics of a system while leaving unchanged its external interface.
- *p* behaviorally refines *q* if the behavior patterns observed for *p* are a structural restriction, with respect to the behavioral model captured by monad B, of those of *q*.
- Any coalgebra ⟨U, p : U → TU⟩ specifies a transition structure over U.
- For extended polynomial functors such a structure may be expressed as a relation →_p⊆ U × U, defined in terms of the *structural membership* relation ∈_T ⊆ U × T U, i.e.,

$$u \longrightarrow_{\rho} u'$$
 iff $u' \in_{\mathsf{T}} \rho u$

- Behavior refinement affects the internal dynamics of a system while leaving unchanged its external interface.
- *p* behaviorally refines *q* if the behavior patterns observed for *p* are a structural restriction, with respect to the behavioral model captured by monad B, of those of *q*.
- Any coalgebra ⟨U, p : U → TU⟩ specifies a transition structure over U.
- For extended polynomial functors such a structure may be expressed as a relation →_p⊆ U × U, defined in terms of the *structural membership* relation ∈_T ⊆ U × T U, i.e.,

$$u \longrightarrow_{p} u'$$
 iff $u' \in_{\mathsf{T}} p u$

- Behavior refinement affects the internal dynamics of a system while leaving unchanged its external interface.
- *p* behaviorally refines *q* if the behavior patterns observed for *p* are a structural restriction, with respect to the behavioral model captured by monad B, of those of *q*.
- Any coalgebra ⟨U, p : U → TU⟩ specifies a transition structure over U.
- For extended polynomial functors such a structure may be expressed as a relation →_p⊆ U × U, defined in terms of the *structural membership* relation ∈_T ⊆ U × T U, i.e.,

$$u \longrightarrow_{\rho} u'$$
 iff $u' \in_{\mathsf{T}} \rho u$



 Structural Membership Relation for extended polynomial functors:

$$\begin{array}{lll} x \in_{\mathsf{Id}} y & \text{iff} & x = y \\ x \in_{\underline{K}} y & \text{iff} & \mathsf{false} \end{array} \\ x \in_{\mathsf{T}_1 \times \mathsf{T}_2} y & \text{iff} & x \in_{\mathsf{T}_1} \pi_1 \ y \ \lor \ x \in_{\mathsf{T}_2} \pi_2 \ y \\ x \in_{\mathsf{T}_1 + \mathsf{T}_2} y & \text{iff} & \begin{cases} y = \iota_1 \ y' & \Rightarrow \ x \in_{\mathsf{T}_1} \ y' \\ y = \iota_2 \ y' & \Rightarrow \ x \in_{\mathsf{T}_2} \ y' \end{array} \\ x \in_{\underline{\mathsf{T}}\underline{\mathsf{K}}} y & \text{iff} & \exists_{k \in K}. \ x \in_{\mathsf{T}} y \ k \\ x \in \mathscr{P}\mathsf{T} \ y & \text{iff} & \exists_{y' \in y'}. \ x \in_{\mathsf{T}} \ y' \end{array}$$



- In data refinement, there is a 'recipe' to identify a refinement situation: look for a *retrieve function* to witness it. I.e., a morphism in the relevant category, from the 'concrete' to the 'abstract' model such that the latter can be *recovered* from the former up to a suitable notion of equivalence, though, typically, not in a unique way.
- In the coalgebraic framework, however, things do not work this way. The reason is obvious: initial states preserving coalgebra morphisms are known to entail bisimilarity. Therefore we have to look for some *weaker* notion of morphism between coalgebras.



- In data refinement, there is a 'recipe' to identify a refinement situation: look for a *retrieve function* to witness it. I.e., a morphism in the relevant category, from the 'concrete' to the 'abstract' model such that the latter can be *recovered* from the former up to a suitable notion of equivalence, though, typically, not in a unique way.
- In the coalgebraic framework, however, things do not work this way. The reason is obvious: initial states preserving coalgebra morphisms are known to entail bisimilarity. Therefore we have to look for some *weaker* notion of morphism between coalgebras.

Introduction	Components as Coalgebras	UML	Logic	Future Directions

• Coalgebra morphism from p to q: B($h \times id$) $\circ \alpha_p = \alpha_q \circ (h \times id)$

• In terms of transitions, the equation can be translated into the requirements:

$$\begin{array}{lll} u \stackrel{\langle i, o \rangle}{\longrightarrow} u' & \Rightarrow & h \ u \stackrel{\langle i, o \rangle}{\longrightarrow} q \ h \ u' \\ h \ u \stackrel{\langle i, o \rangle}{\longrightarrow} v' & \Rightarrow & \exists_{u' \in U}. \ u \stackrel{\langle i, o \rangle}{\longrightarrow} u' \ \land \ u' = h \ v' \end{array}$$

which jointly state that, not only *p* dynamics is *preserved* by *h*, but also *q* dynamics is *reflected* back over the same *h*.

Introduction	Components as Coalgebras	UML	Logic	Future Directions

• Coalgebra morphism from *p* to *q*:

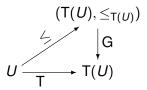
$$\mathsf{B}(h\times\mathsf{id})\circ\alpha_p\ =\ \alpha_q\circ(h\times\mathsf{id})$$

 In terms of transitions, the equation can be translated into the requirements:

$$\begin{array}{rcl} u \stackrel{\langle i, o \rangle}{\longrightarrow} u' & \Rightarrow & h \ u \stackrel{\langle i, o \rangle}{\longrightarrow} q \ h \ u' \\ h \ u \stackrel{\langle i, o \rangle}{\longrightarrow} v' & \Rightarrow & \exists_{u' \in U} \cdot \ u \stackrel{\langle i, o \rangle}{\longrightarrow} u' \ \land \ u' = h \ v' \end{array}$$

which jointly state that, not only *p* dynamics is *preserved* by *h*, but also *q* dynamics is *reflected* back over the same *h*.

Given a **Set** endofunctor T, an order \leq on T is defined as a functor \leq from **Set** to **PreOrder** (concretely, mapping every set *U* into a collection of preorders $\leq_{T(U)}$) making the following diagram to commute:



where G is the forgetful functor which forgets the preorder structure for every preordered set and gives its underlying set.



Let T be an extended polynomial functor on Set and consider two T-coalgebras c = (U, α : U → TU) and a = (V, β : V → TV). A *forward* morphism h : c → a with respect to a refinement preorder ≤, is a function from U to V such that

 $\mathsf{T} h \circ \alpha \leq \beta \circ h$

- Dually, *h* is called a *backward* morphism if $\beta \circ h \leq Th \circ \alpha$.
- For coalgebras *c* and *a*, *c* is a *behavior refinement* of *a*, written *c* ⊆_B *a*, if there exist coalgebras *p* and *q* such that *c* ~ *p*, *a* ~ *q* and there exists a (initial state preserving) forward morphism from *p* to *q*.



Let T be an extended polynomial functor on Set and consider two T-coalgebras c = (U, α : U → TU) and a = (V, β : V → TV). A *forward* morphism h : c → a with respect to a refinement preorder ≤, is a function from U to V such that

$$\mathsf{T}\boldsymbol{h}\circ\alpha\leq\beta\circ\boldsymbol{h}$$

- Dually, *h* is called a *backward* morphism if $\beta \circ h \leq Th \circ \alpha$.
- For coalgebras *c* and *a*, *c* is a *behavior refinement* of *a*, written *c* ⊑_B *a*, if there exist coalgebras *p* and *q* such that *c* ~ *p*, *a* ~ *q* and there exists a (initial state preserving) forward morphism from *p* to *q*.

Behavioral Refinement

The exact meaning of a refinement assertion $c \sqsubseteq_B a$ depends on the refinement preorder \leq adopted. For example, we can define the preorder for extended polynomial functor T by induction as follows:

$x \subseteq_{\mathit{ld}} y$	iff	x = y
$x\subseteq_{K} y$	iff	$x =_{K} y$
$x \subseteq_{T_1 \times T_2} y$	iff	$\pi_1 \mathbf{X} \subseteq_{T_1} \pi_1 \mathbf{y} \ \land \ \pi_2 \mathbf{X} \subseteq_{T_2} \pi_2 \mathbf{y}$
$x \subseteq_{T_1+T_2} y$	iff	$\begin{cases} x = \iota_1 x' \land y = \iota_1 y' \Rightarrow x' \subseteq_{T_1} y' \\ x = \iota_2 x' \land y = \iota_2 y' \Rightarrow x' \subseteq_{T_2} y' \end{cases}$
$\pmb{x}\subseteq_{T^{K}}\pmb{y}$	iff	$orall k \in \mathcal{K}. \ x(k) \subseteq_{T} y(k)$
$x \subseteq_{\mathscr{P}T} y$	iff	$\forall \boldsymbol{e} \in \boldsymbol{x}. \ \exists \boldsymbol{e}' \in \boldsymbol{y}. \ \boldsymbol{e} \subseteq_{T} \boldsymbol{e}'$

- UML is "a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system".
- In practice, it stands for a collection of inter-related, semiformal design notations for software development, providing a unified notation, expressive and widely adopted.
- It lacks a rigourous and consensual semantic definition leading, therefore, to weak effective support to the design of complex systems and, often, to conflicting support tools.



- UML is "a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system".
- In practice, it stands for a collection of inter-related, semiformal design notations for software development, providing a unified notation, expressive and widely adopted.
- It lacks a rigourous and consensual semantic definition leading, therefore, to weak effective support to the design of complex systems and, often, to conflicting support tools.



- UML is "a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a softwareintensive system".
- In practice, it stands for a collection of inter-related, semiformal design notations for software development, providing a unified notation, expressive and widely adopted.
- It lacks a rigourous and consensual semantic definition leading, therefore, to weak effective support to the design of complex systems and, often, to conflicting support tools.



- We introduced a generic coalgebraic semantic framework for different models in UML, including class diagrams, use cases, statecharts and sequence diagrams, where the semantics of different kinds of models are given as coalgebras.
- Notions of bisimulation and refinement capture observational equivalence and simulation preorders, respectively, and form the basis of a whole discipline of reasoning and transforming UML designs.



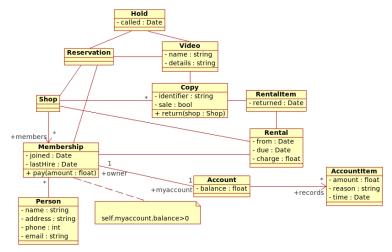
- We introduced a generic coalgebraic semantic framework for different models in UML, including class diagrams, use cases, statecharts and sequence diagrams, where the semantics of different kinds of models are given as coalgebras.
- Notions of bisimulation and refinement capture observational equivalence and simulation preorders, respectively, and form the basis of a whole discipline of reasoning and transforming UML designs.



- In UML a *class diagram* captures the static structure of a system, as a set of classes and relationships, called *associations*, between them.
- Classes may be further annotated with *constraints*, i.e., properties that must hold for every object in the class along its lifetime.
- We concentrate here in class declarations. The aim of a class declaration is introduce a signature of attributes and methods.

Class Diagrams

• Consider the simplified model of a video renting e-business:



Class Diagrams

• Consider class **Membership** in the previous example. It introduces two attributes and a method over a state space, identified by variable *U* below, which is made observable exactly (and uniquely) by the attributes and methods it declares:

> joined : $U \longrightarrow Date$ lastHire : $U \longrightarrow Date$ pay : $U \times \mathbb{R} \longrightarrow U$

These three declarations can be grouped in one through a *split* construction

 $\langle \text{joined}, \text{lastHire}, \overline{\text{pay}} \rangle : U \longrightarrow Date \times Date \times U^{\mathbb{R}}$

which is a *coalgebra* for functor $T X = Date \times Date \times X^{\mathbb{R}}$.

Class Diagrams

• In general, the semantics [[c]] of a class c is given by a specification of a coalgebra

$$\langle \operatorname{at}, \operatorname{\overline{md}} \rangle : U \longrightarrow A \times (O \times U)^{I}$$

where A is the attribute domain, and each method accepts a parameter, of type I, and delivers both a state change and an output value, of type O. I.e., a coalgebra for functor

$$\mathsf{T}: X \longrightarrow A \times (O \times X)^{l}$$

Typically, *I* and *O* are *sum* types, aggregating the inputoutput parameters of each declared method. On its turn, *A* is usually a *product* type joining all attribute outputs in a way which emphasises that each of them is available independent of the others, and therefore always able to be accessed in parallel.



• More generally, as methods are typically implemented by *partial functions* or even by arbitrary *relations*, this definition should be generalized to

$$\langle \mathsf{at}, \overline{\mathsf{md}} \rangle : U \longrightarrow A \times \mathsf{B}(O \times U)^{T}$$

where B is a strong monad capturing some sort of behavioral effect.

Class Diagrams

- A UML class diagram introduces a number of class specifications which types the object population of any corresponding model implementation. Typically, different ways of putting classes together in a class diagram correspond to different operators between the T-coalgebras.
- In particular, one may consider a form of *parallel* aggregation, denoted by ⊠, in which methods in both classes can be called simultaneously (as they always act upon disjoint state spaces), and a form of *interleaving*, denoted by ⊞, which offers a choice of which class to call.
- Note that in both cases, attributes are always available to be observed, and therefore are composed in a multiplicative context. Initial conditions are joined by logical conjunction.



 Constraints are typically attached to class specifications and their semantic effect is to constraint what coalgebras count as valid implementation for the class. Such is the case, for example, of constraint

balance > 0

attached to class Membership in our example.

Class Diagrams

- Associations can also be interpreted as constraints, with respect to a fragment of the diagram containing the two associated classes.
- For this, one has to assume that the state space of each class has a component recording the collection of live instances.
- An *association* becomes a constraint over such components of the (joint) state space.
- For example a 'one-to-one' association corresponds to a predicate asserting the existence of an injective function relating the collection of instances of each class.
- Similarly, a 'one-to-many' association corresponds to a relation whose kernel is the identity, i.e., a total relation whose converse is simple.

Class Diagrams

- In general, constraints and associations are predicates which are supposed to be preserved along the system life-time.
- Formally, they are incorporated in the semantics as *invariants*. Such predicates, once encoded as coreflexives, i.e., fragments of the identity, according to

$$y \Phi_P x \equiv y = x \wedge P x$$

can be specified as $c \cdot \Phi_P \subseteq T \Phi_P \cdot c$.

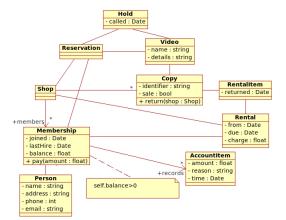
 When reasoning about diagram transformations, such as refactoring, constraints entail for *proof obligations*. For example,

 $\llbracket balance > 0 \rrbracket = \\ \llbracket Membership \rrbracket \cdot \Phi_{balance > 0} \subset T \Phi_{balance > 0} \cdot \llbracket Membership \rrbracket$

needs to be discarded whenever justifying a refactoring involving class **Membership**.

Class Diagrams

- Consider the *Inline Class Refactoring* pattern: Inline class refactoring allows two classes to be merged together provided one of them has no methods available.
- The previous example can be transformed into the following diagram:



Class Diagrams

 classes Membership and Account are replaced by a new class Membership' whose semantics is a new coalgebra over the state space of [[Membership]] to which a new attribute balance is added.

 $[\![Membership']\!] = \langle \langle at_{Membership}, at_{Account} \rangle, \overline{md}_{Membership} \rangle$

 Assuming the remaining part of the diagram remains unchanged, clearly the new class [[Membership']] and the relevant fragment of the original class diagram, i.e.,

[Membership] ⊠ [Account]

are not bisimilar.

• But we can prove that the inline refactoring is acturally a refinement between the two diagrams.



- Sun Meng, Zhang Naixiao and Luis Barbosa. On Semantics and Refinement of UML Statecharts: A Coalgebraic View. In Proceedings of SEFM'04, pages 164-173, IEEE Computer Society, 2004.
- Sun Meng and Luis Barbosa. A Coalgebraic Semantic Framework for Reasoning about UML Sequence Diagrams. In Proceedings of QSIC'08. IEEE Computer Society, 2008.
- Sun Meng and Luis Barbosa. A Coalgebraic Semantic Framework for Reasoning about Interaction Designs. In Kevin Lano eds. UML Semantics and its Applications. pages 249-280, Wiley, 2009.



- Coalgebra offers tools that apply uniformly to a large class of systems.
- An obvious question from this perspective is whether we can deal with logics for coalgebras in a uniform way.
- This question is of interest from a computer science point of view because coalgebras are systems and logics are specification languages.



- As an example, consider the coalgebraic logic invented by Lawrence Moss, for which the syntax and semantics working in a uniform way for all signatures Σ : Set → Set.
- Formulas of the logic are invariants under behavioral equivalence and the logic is reasonably expressive.
- Here "reasonably expressive" means by requiring that admitting infinite conjunctions, the logic should be able to characterize processes (elements of coalgebras) up to behavioral equivalence.
- The aim is to find a language L_Σ and for each Σ-coalgebra (X, ψ) a relation ⊨_Σ⊂ X × L_Σ satisfying the above requirements.

UML

Coalgebra and Logic ...

- The starting point is that signatures are functors on Set and may hence also be applied to sets of formulas L_Σ and relations ⊨_Σ.
- Functors Σ on Set are extended to functors on the category of classes SET via ΣK = U{ΣX : X ⊂ K, X a set} for classes K. Moreover, Σ is assumed to weakly preserve pullbacks.
- The syntax of coalgebraic logic:

Definition

 \mathcal{L}_{Σ} is defined to be the least class satisfying:

```
egin{array}{lll} \Phi \subset \mathcal{L}_{\Sigma}, \Phi 	ext{ a set } \Rightarrow & \wedge \Phi \in \mathcal{L}_{\Sigma} \ & \phi \in \Sigma(\mathcal{L}_{\Sigma}) \Rightarrow & \phi \in \mathcal{L}_{\Sigma} \end{array}
```

That is, \mathcal{L}_{Σ} is the initial algebra wrt the functor $\mathscr{P} + \Sigma$.

UML

Coalgebra and Logic ...

- The starting point is that signatures are functors on Set and may hence also be applied to sets of formulas L_Σ and relations ⊨_Σ.
- Functors Σ on Set are extended to functors on the category of classes SET via ΣK = U{ΣX : X ⊂ K, X a set} for classes K. Moreover, Σ is assumed to weakly preserve pullbacks.
- The syntax of coalgebraic logic:

Definition

 \mathcal{L}_{Σ} is defined to be the least class satisfying:

$$egin{array}{lll} \Phi \subset \mathcal{L}_{\Sigma}, \Phi ext{ a set } \Rightarrow & \wedge \Phi \in \mathcal{L}_{\Sigma} \ & \phi \in \Sigma(\mathcal{L}_{\Sigma}) \Rightarrow & \phi \in \mathcal{L}_{\Sigma} \end{array}$$

That is, \mathcal{L}_{Σ} is the initial algebra wrt the functor $\mathscr{P} + \Sigma$.

• The semantics of coalgebraic logic goes as follows:

Definition

Given a coalgebra (X, ψ) define $\models_{\Sigma} \subset X \times \mathcal{L}_{\Sigma}$ as the least relation such that (let $x \in X$):

$$\begin{aligned} x \models_{\Sigma} \phi \text{ for all } \phi \in \Phi, \Phi \subset \mathcal{L}_{\Sigma}, \Phi \text{ a set } \Rightarrow x \models_{\Sigma} \bigwedge \Phi \\ \exists w \in \Sigma(\models_{\Sigma}) s.t. \Sigma \pi_{1}(w) = \psi(x), \Sigma \pi_{2}(w) = \phi \Rightarrow x \models_{\Sigma} \phi \end{aligned}$$

where π_1, π_2 denotes the projections from the product $X \times \mathcal{L}_{\Sigma}$ to its components.

• The following theorem shows that coalgebraic logic reflects precisely the notion of behavioral equivalence:

Theorem

Let $\boldsymbol{\Sigma}$ be a functor on \boldsymbol{Set} which weakly preserving pullbacks. Then

- 1. Formulas of \mathcal{L}_{Σ} are invariant under behavioral equivalence and
- **2.** For each coalgebra (X, ψ) and each $x \in X$ there is a formula $\phi_x \in \mathcal{L}_{\Sigma}$ such that for all coalgebras (X', ψ') and all $x' \in X'$,

 $x' \models_{\Sigma} \phi_x$ iff x, x' behaviorally equivalent

For more details about the coalgebraic logic:

Lawrence Moss. Coalgebraic logic. *Annals of Pure and Applied Logic*. 96: 277-317, 1999.

UML

Coalgebra and Logic ...

In fact, I am not an expert in this area ...

But there are many references:

- Yde Venema. Algebras and coalgebras. In *Handbook of Modal Logic*, pages 331-426, Elsevier, 2006.
- Alexander Kurz. Coalgebras and their Logics. *SIGACT News*, vol. 37, pages 57-77, 2006.
- Bart Jacobs. The temporal logic of coalgebras via Galois algebras. *Mathematical Structures in Computer Science*, vol. 12(6), pages 875-903, 2002.

In fact, I am not an expert in this area ...

But there are many references:

- Yde Venema. Algebras and coalgebras. In *Handbook of Modal Logic*, pages 331-426, Elsevier, 2006.
- Alexander Kurz. Coalgebras and their Logics. SIGACT News, vol. 37, pages 57-77, 2006.
- Bart Jacobs. The temporal logic of coalgebras via Galois algebras. *Mathematical Structures in Computer Science*, vol. 12(6), pages 875-903, 2002.

- A coalgebraic perspective on evolving state-based systems is of interest to logicians, because ...
 - Uniform treatment of different types of systems. For example, one can establish that satisfiability of coalgebraic logic is in PSPACE and that complete coalgebraic logics have the finite model property. The uniformity of the metatheory might well translate into software tools that are easier to design, maintain, and to implement.
 - Modularity. Different functors can be combined using composition of functors, product, coproduct, etc. Theorems and algorithms for basic types can then be lifted to arbitrarily complex combinations.
 - One-step analysis. Coalgebraic analysis of dynamic systems is particularly successful where the class of all complete behaviors is determined by the possible one-step behaviors. This is the basis of coinduction. It also plays an important role in applications to automata theory.

- A coalgebraic perspective on evolving state-based systems is of interest to logicians, because ...
 - Uniform treatment of different types of systems. For example, one can establish that satisfiability of coalgebraic logic is in PSPACE and that complete coalgebraic logics have the finite model property. The uniformity of the metatheory might well translate into software tools that are easier to design, maintain, and to implement.
 - Modularity. Different functors can be combined using composition of functors, product, coproduct, etc. Theorems and algorithms for basic types can then be lifted to arbitrarily complex combinations.
 - One-step analysis. Coalgebraic analysis of dynamic systems is particularly successful where the class of all complete behaviors is determined by the possible one-step behaviors. This is the basis of coinduction. It also plays an important role in applications to automata theory.

- A coalgebraic perspective on evolving state-based systems is of interest to logicians, because ...
 - Uniform treatment of different types of systems. For example, one can establish that satisfiability of coalgebraic logic is in PSPACE and that complete coalgebraic logics have the finite model property. The uniformity of the metatheory might well translate into software tools that are easier to design, maintain, and to implement.
 - Modularity. Different functors can be combined using composition of functors, product, coproduct, etc. Theorems and algorithms for basic types can then be lifted to arbitrarily complex combinations.
 - One-step analysis. Coalgebraic analysis of dynamic systems is particularly successful where the class of all complete behaviors is determined by the possible one-step behaviors. This is the basis of coinduction. It also plays an important role in applications to automata theory.

- A coalgebraic perspective on evolving state-based systems is of interest to logicians, because ...
 - Uniform treatment of different types of systems. For example, one can establish that satisfiability of coalgebraic logic is in PSPACE and that complete coalgebraic logics have the finite model property. The uniformity of the metatheory might well translate into software tools that are easier to design, maintain, and to implement.
 - Modularity. Different functors can be combined using composition of functors, product, coproduct, etc. Theorems and algorithms for basic types can then be lifted to arbitrarily complex combinations.
 - One-step analysis. Coalgebraic analysis of dynamic systems is particularly successful where the class of all complete behaviors is determined by the possible one-step behaviors. This is the basis of coinduction. It also plays an important role in applications to automata theory.

• .



- Coalgebras have been successfully applied to many areas in computer science.
- Applications in mathematics and logic have been developed much less than applications to computer science, but coalgebra as an area naturally overlaps with Universal Algebra, Modal Logic, Domain Theory and Category Theory.
- We believe that this presents many opportunities for exciting future research in different areas.

Introduction	Components as Coalgebras	UML	Logic	Future Directions

Thank you!