SENG 697
Agent-based Software Engineering

Session 2 : Methodologies for agent-based analysis and design

Behrouz Far
Schulich School of Engineering, University of Calgary
far@ucalgary.ca
http://www.enel.ucalgary.ca/People/far/
Course Curriculum

Overview of agent-based SE

Methodologies for agent-based analysis and design

Agent communication & knowledge sharing

Agent-based System Architecture & Organization

FIPA: Foundation for Intelligent Physical Agents

Principles of Object Technology

Other topics: Agent Interaction, Infrastructure, APIs, Performance metrics, Learning, Self-organizing systems etc.

far@ucalgary.ca
Multi Agent system Development

What do we need to develop an agent-based system?
Agent-based development **methodology**
- How to analysis and design a multi-agent system?
  - e.g., GAIA, Tropos, Prometheus, etc.

Agent-based development **platform**
- How to implement a multi-agent system?
  - e.g., FIPA-OS, Jade, Jive, Aglets, etc.
Evolution of software engineering paradigms:
- Assembly languages
- Procedural and structured programming
- Object Oriented programming
- Component-ware
- Design patterns
- Software architectures
- Software Agents

Languages that have their conceptual basis determined by machine architecture

Increase of Complexity

Languages that have their key abstractions rooted in the problem domain
Why We Need This?

- The authors have devised a methodology that has been specifically tailored to the analysis and design of agent-based systems.
- Why we need this?
Why We Need This? (contd.)

- The way Object Oriented systems are designed
- Objects and their “pre-defined” interactions are designed
- Behaviour is the result of objects interactions
Why We Need This? (contd.)

- The way agent-based systems are designed
- Agents, their “roles” in the society, “goals” and “possible” ways of interactions are defined.
- Behaviour is the result of agents interactions
Why We Need This? (contd.)

- Existing software development methodologies are rather unsuitable for agent based software engineering tasks.
- There is a fundamental mismatch between the concepts used by object-oriented developers and the agent-oriented view. They fail to adequately capture:
  - Agents’ flexible, autonomous (problem-solving) behaviour
  - Richness of agent interactions
  - Complexity of an agent system organizational structures
Contents

- **Session 2 : Methodologies for agent-based analysis and design**
  - A Methodology for Agent-Oriented Analysis and Design (GAIA)
  - Tropos
  - The MASSIVE Development Method
  - Comparison of development techniques
A Methodology for Agent-Oriented Analysis and Design: GAIA

Michael Wooldridge
Nicholas R. Jennings
David Kinny
Abstract

- GAIA is a methodology for agent-oriented analysis and design. ← detailed analysis

- GAIA methodology is concerned with how a society of agents cooperate to realize the system level goals, and what is required of each individual agent in order to do this.

- The methodology is general. It is applicable to a wide range of multi-agent systems. It deals with both the macro-level (societal) and the micro-level (agent) aspects of systems.

- The methodology is founded on the view of a system as a computational organization consisting of various interacting roles.
Introduction

- Progress in software engineering over the past two decades has primarily been made through the development of increasingly powerful and natural abstractions with which to model and develop complex systems.

- Procedural abstraction, abstract data types, and, most recently, objects, are all examples of such abstractions.

- It is the authors’ belief that agents represent a similar advance in abstraction: they may be used by software developers to more naturally understand, model, and develop an important class of complex distributed systems.
Agents are coarse-grained computational systems, each making use of significant computational resources (think of each agent as having the resources of a UNIX process.)

It is assumed that the goal is to obtain a system that maximizes some global quality measure, but which may be suboptimal from the point of view of the system components.

Agents are heterogeneous, in that different agents may be implemented using different programming languages and techniques.

The overall system contains a comparatively small number of agents (usually less than 100).
Conceptual Framework

- The methodology is intended to allow an analyst to go systematically from a statement of requirements to a design that is sufficiently detailed that it can be implemented directly.

- In applying the methodology, the analyst moves from abstract to increasingly concrete concepts.

- Each successive move introduces greater implementation bias, and shrinks the space of possible systems that could be implemented to satisfy the original requirements statement.

- The methodology encourages a developer to think of building agent based systems as a process of organizational design.
Analysis Concepts

System

Role(s)

By means of?
- Permission

What?
- Responsibility

How?
- Protocol

Liveness property
(goal, utility)

Safety property
(guards)

far@ucalgary.ca 15
Abstract & Concrete Concepts

The concepts used by GAIA are split in two categories:

- **Abstract**: used in the analysis phase, that don’t necessarily have a correspondent in the implementation of the system
- **Concrete**: concepts that typically have a correspondent in run-time system.

<table>
<thead>
<tr>
<th>Abstract Concepts</th>
<th>Concrete Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roles</td>
<td>Agent Types</td>
</tr>
<tr>
<td>Permissions</td>
<td>Services</td>
</tr>
<tr>
<td>Responsibilities</td>
<td>Acquaintances</td>
</tr>
<tr>
<td>Liveness properties</td>
<td></td>
</tr>
<tr>
<td>Safety properties</td>
<td></td>
</tr>
<tr>
<td>Protocols</td>
<td></td>
</tr>
<tr>
<td>Activities</td>
<td></td>
</tr>
</tbody>
</table>
1. **System = Society of Agents**

- The most abstract entity in the concept hierarchy is the *system* which is equivalent to society or organization of agents. That is, we think of an agent-based system as an *artificial society* or *organization*.

- The idea of a system as a society is useful when thinking about the next level in the concept hierarchy: *roles*.
2. Role

- It may seem strange to think of a computer system as being defined by a set of *roles*, but the idea is quite natural when adopting an organizational view of the world.
- Consider a human organization such as a typical company. The company has roles such as president, vice president, etc.
- The roles will be instantiated with actual individuals: i.e., individuals take on the role of president, vice president, etc.
- The instantiation is not necessarily static. E.g., throughout the company lifetime, many individuals may take on the role of company president.
- A role is defined by three attributes: *responsibilities*, *permissions*, and *protocols*.
3. Responsibilities (What?)

- Responsibilities determine functionality and are the key attribute associated with a role.
- **Example:** a responsibility associated with the role of company president might be calling the shareholders meeting every year.
- Responsibilities are divided into two types: *liveness properties* and *safety properties*. 
4. Liveness (Goal or Utility)

- **Liveness properties** describe what must be achieved by a role under certain environmental conditions.

- **Example:** increase profit for “seller” role. Minimize the price while getting the best quality for “buyer” role.
5. Safety (Guards)

- **Safety properties** impose a number of constraints on the behavioural states of a role (i.e., that an acceptable state of affairs is maintained across all states of execution).

- **Example:** Ensure the liabilities of the company are always less than assets.
6. Permissions = Rights

- In order to realise responsibilities, a role is usually associated with a set of permissions.
- Permissions are the rights associated with a role. The permissions of a role thus identify the resources that are available to that role in order to realize its responsibilities.
- In typical systems, permissions tend to be information resources.

**Examples:**
- Ability to retrieve a particular information; modify or delete another piece of information.
- Ability to generate information.
7. Protocols (How?)

- A role is also identified with a number of protocols, which define the way that a role can interact with the other roles in the system.

- **Example:**
  - A *seller* role might have the protocols *Dutch auction* and *English auction* associated with it.
  - A *hiring* role of the president of a company is to conform to the hiring protocol (policies?).
Roles: how to find them?

- Think of possible scenario of system behavior and identify actors in them
- Think of possible must do and must not do actions
- Think of what is needed to do a task and what should be avoided
- Think of possible interactions
How to Find Roles?

Roles

Actors

Use cases

Supplementary Specifications

Glossary

Use-Case Model

Use-Case Realization (identified)

Software Architecture Document

Views:
- Conceptual view
- Module view
- Code view
- Execution view
### Case of Agile Development

**Main source:** User stories

**Example:**

<table>
<thead>
<tr>
<th>Customer</th>
<th>BSC Clerk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>After inspecting/completing a request, can mark the status as held/complete. If &quot;hold&quot;, can enter the reason. If &quot;complete&quot;, must have all documents attached.</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>4.0 hours</td>
</tr>
</tbody>
</table>
| **Acceptance Tests** | 1. Mark as complete with required documents not included. (fail)  
2. Mark as complete with required documents included. (pass)  
3. Mark as held. (pass) |
Analysis & Design Models

Requirements Statement

Roles Model

Interactions Model

Agent Model

Services Model

Acquaintances Model

Analysis

Design aka. Detailed analysis

far@ucalgary.ca 27
1) Analysis Phase

- The objective of the analysis phase is to develop an understanding of the system and its structure as a *multi-agent organization*. This understanding is captured in the system organization.

- *Organization* is viewed as a collection of roles, that stand in certain relationships to one another, and that take part in systematic, institutionalised patterns of interactions with other roles.

- To define an organization, it suffices to define the roles in the organization, how these roles relate to one another, and how a role can interact with other roles.

- The analysis phase is comprised of two models:
  - *Role model*
  - *Interactions model*
Analysis: Roles Model

- The *roles model* identifies the key roles in the system. Here a role can be viewed as an abstract description of an entity’s expected function.
- Roles are characterised by two types of attribute:
  - **The permissions/rights associated with the role**
    - A role will have associated with it certain permissions, relating to the type and the amount of resources that can be exploited when carrying out the role. These aspects are captured in an attribute known as the role's permissions.
  - **The responsibilities of the role**
    - A role is created in order to do something. That is, a role has a certain functionality. This functionality is represented by an attribute known as the role's responsibilities.
Analysis: Permissions

- The *permissions* associated with a role have two aspects:
  - They identify the resources that can legitimately be used to carry out the role, intuitively, they say what can be spent while carrying out the role;
  - They state the resource limits within which the role executor must operate, intuitively, they say what can be spent while carrying out the role.

- In this method, in order to carry out a role, an agent will typically be able to access certain information (or other resources).

- Some roles might generate information; others may need to access a piece of information but not modify it, and others may need to modify the information.
The functionality of a role is defined by its **responsibilities**. These responsibilities can be divided into two categories: liveness responsibilities and safety conditions.

**Liveness responsibilities** are so called because they tend to say that something will be done and hence that the agent carrying out the role is still alive.

- Liveness responsibilities tend to follow certain patterns.
- **Example:** The guaranteed response type of achievement goal has the form “request is always followed by a response”.

- Liveness expressions define the potential execution trajectories through the various activities and interactions (i.e., over the protocols) associated with the role.
In many cases, it is insufficient simply to specify the liveness responsibilities of a role. This is because an agent, carrying out a role, will be required to maintain certain invariants while executing.

**Example:** We might require that a particular agent taking part in an electronic commerce application never spends more money than it has been allocated.

These invariants are called *safety conditions*, because they usually relate to the absence of some undesirable condition arising.

Safety requirements are specified by means of a list of predicates. These predicates are typically expressed over the variables listed in a role’s permissions attribute.
GAIA: Analysis Process

1. Identify the **roles** in the system.
   - **Output:** A prototypical roles model, a list of the key roles that occur in the system, each with an informal, unelaborated.

2. For each role, identify and document the associated protocols. Protocols are the patterns of *interaction* that occur in the system between the various roles.
   - **Output:** An interaction model, which captures the recurring patterns of inter-role interaction.

3. Using the protocol model as a basis, elaborate the roles model.
   - **Output:** A fully elaborated roles model, which documents the key roles occurring in the system, their permissions and responsibilities, and the protocols in which they take part.

4. Iterate stages (1)- (3).
Multi-Agent Dispatch System (MADS) provides Road-side assistance system for a fleet of vehicles.
Example: MADS System

- **Role: Geocode**
  **Description:** Wraps the MapPoint geocoding Web Services and it queries MapPoint for geocodes.

- **Role: RenderMap**
  **Description:** It wraps the MapPoint map rendering Web Services and it queries MapPoint for maps.

- **Role: Route**
  **Description:** It wraps the MapPoint routing Web Services and it queries MapPoint for routes.

- **Role: PersonalAssistant**
  **Description:** It acts on behalf of a system user. It gets the geocode information, maps and routes, and it calculates distances between the breakdown and available service providers.
Example (cont’d)

Permissions for RenderMap:

- reads
  - MapPointWebServices
  - newMap // true or false

- changes
  - notifyUser // true or false
Example (cont’d)

Liveness Property:

- When the newMap is true, get a new map.
- When the new map is ready, notify the user.

RENDERMAP = (RenderMap)
RENDERMAP = RequestMap.QueryMapPoint.RespondMap
Safety Property:

- Safety properties are specified using a list of predicates in relationship with the variables listed in the role’s permissions attribute.

- For the *RenderMap* role, a successful connection with MapPoint Web Services is the safety property.
GAIA roles model is composed of a set of role schemata for each role in the system with the following template:

<table>
<thead>
<tr>
<th>Role Schema</th>
<th>name of role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>short description of the role</td>
</tr>
<tr>
<td>Protocols and Activities</td>
<td>protocols and activities in which the role plays a part</td>
</tr>
<tr>
<td>Permissions</td>
<td>rights associated with the role</td>
</tr>
<tr>
<td>Responsibilities</td>
<td>liveness responsibilities/properties</td>
</tr>
<tr>
<td></td>
<td>safety responsibilities/properties</td>
</tr>
</tbody>
</table>
Example: **RenderMap**

<table>
<thead>
<tr>
<th>Role Schema</th>
<th><strong>RenderMap</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>It wraps the MapPoint map rendering Web Services and it queries MapPoint for maps.</td>
</tr>
<tr>
<td><strong>Protocols and Activities</strong></td>
<td><strong>QueryMapPoint</strong>, <strong>RequestMap</strong>, <strong>RespondMap</strong></td>
</tr>
<tr>
<td><strong>Permissions</strong></td>
<td>reads MapPoint Web Services</td>
</tr>
<tr>
<td><strong>Responsibilities</strong></td>
<td><strong>Liveness</strong></td>
</tr>
<tr>
<td></td>
<td>RENDERMAP = (RenderMap)</td>
</tr>
<tr>
<td></td>
<td>RENDERMAP = RequestMap.QueryMapPoint.RespondMap</td>
</tr>
<tr>
<td></td>
<td><strong>Safety</strong> a successful connection with MapPoint Web Services</td>
</tr>
</tbody>
</table>
There are dependencies and relationships between the various roles in a multi-agent organisation. Interactions need to be captured and represented in the analysis phase.

This model consists of a set of protocol definitions (mechanisms), one for each type of inter-role interaction.

A protocol can be viewed as a pattern of interaction that has been formally defined and abstracted away from any particular sequence of execution steps.

A single protocol definition will give rise to a number of message interchanges in the run time system.

Example: English auction protocol. This involves multiple roles (sellers and bidders) and many potential patterns of interchange (price announcements and corresponding bids). However at the analysis stage, such precise instantiation details are not necessary.
Analysis: Interaction Model

- Protocol definitions consist of the following set of attributes:
  - **Purpose**: brief description of the nature of the interaction (e.g. Information request, schedule activity and assign task);
  - **Initiator**: the role(s) responsible for starting the interaction;
  - **Responder**: the role(s) with which the initiator interacts;
  - **Inputs**: information used by the role initiator while enacting the protocol;
  - **Outputs**: information supplied by/to the protocol responder during interaction;
  - **Processing**: brief description of any processing the protocol initiator performs during the course of the interaction.
### Example: Interaction Model

<table>
<thead>
<tr>
<th><strong>Protocol</strong></th>
<th><strong>RequestGeocode</strong></th>
<th><strong>RequestMap</strong></th>
<th><strong>RequestRoute</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose/parameters</strong></td>
<td>Request to geocode a location. The response includes the Latitude and Longitude of the location</td>
<td>Request a map containing a list of location. The response include the URL to the map that displays the locations</td>
<td>Request a route between two locations. The response includes the driving direction between the locations</td>
</tr>
<tr>
<td><strong>Initiator(s)</strong></td>
<td>PersonalAssistant</td>
<td>PersonalAssistant</td>
<td>PersonalAssistant</td>
</tr>
<tr>
<td><strong>Receiver(s)</strong></td>
<td>Geocode</td>
<td>RenderMap</td>
<td>Route</td>
</tr>
<tr>
<td><strong>Responding Protocol</strong></td>
<td>RespondGeocode</td>
<td>RespondMap</td>
<td>RespondRoute</td>
</tr>
</tbody>
</table>
2) Design Phase

- The aim of the design phase is to transform the analysis models into a lower level of abstraction that traditional design techniques (e.g. object-oriented techniques) may be applied.

- How an agent realizes its services and how it is implemented is beyond the scope of the methodology, and will depend on the particular application domain.
Analysis & Design Models

- Requirements Statement
  - Roles Model
    - Agent Model
    - Services Model
  - Interactions Model
    - Acquaintances Model

Analysis → Design → detailed analysis
The design process involves generating three models.

- **Agent model:** identifies the agent types that will make up the system, and the agent instances that will be instantiated from these types.

- **Service (function or operation) model:** identifies the main services that will be associated with each agent type.

- **Acquaintance (collaboration) model:** documents the acquaintances for each agent type.
Design: Agent Model

- The purpose of the **agent model** is to document the various agent types that will be used in the system, and the agent instances that will realise these agent types at run-time.
- An **agent type** is composed of a set of **agent roles**.
- There may in fact be a one-to-one correspondence between roles and agent types. However, this need not be the case. A designer can choose to package a number of closely related roles in the same agent type for the purposes of convenience.
- The agent model is defined using a simple tree, in which root nodes correspond to agent types and leaf nodes correspond to roles, (as defined in the roles model).
Agent Model

<table>
<thead>
<tr>
<th>Qualifier</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>there will be exactly ( n ) instances</td>
</tr>
<tr>
<td>m ( \ldots ) n</td>
<td>there will be between ( m ) and ( n ) instances</td>
</tr>
<tr>
<td>*</td>
<td>there will be 0 or more instances</td>
</tr>
<tr>
<td>+</td>
<td>there will be 1 or more instances</td>
</tr>
</tbody>
</table>

- AgentType1
  - Role1
  - Role2

- AgentType2
  - Role3

- Between 2 and 4 instances

- There will be exactly 3 instances

- There will be 0 or more instances

- There will be 1 or more instances

far@ucalgary.ca
Example: Agent Model

Legend
- Agent types
- Agent roles

Graph:
- Geocode
- RenderMap
- Route
- PersonalAssistant
- Far@ucalgary.ca
The aim of the *services model* is to identify the services associated with each agent role, and to specify the main properties of these services.

Service means a function or an operation of the agent. A service is a single, coherent block of activity that an agent will engage in.

In OO terms, a service would correspond to an operation (a method); however, the services may not be available for other agents in the same way that an object’s methods are available for another object to invoke.
For each service that may be performed by an agent, it is necessary to document its properties. Specifically, we must identify the inputs, outputs, pre-conditions, and post-conditions of each service.

Inputs and outputs to services will be derived from the protocols model.

Pre-conditions and post-conditions represent constraints on services. These are derived from the safety properties of a role. By definition, each role will be associated with at least one service.

The services that an agent will perform are derived from the list of protocols and responsibilities associated with a role, and in particular, from the liveness definition of a role.
## Example: Service Model

<table>
<thead>
<tr>
<th>Service</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Pre-conditions</th>
<th>Post-conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geocode</td>
<td>Address</td>
<td>Latitude and Longitude of the Address</td>
<td>A personalized assistant agent is created and associated with the user</td>
<td>User enters an address</td>
</tr>
<tr>
<td>RenderMap</td>
<td>List of geocodes</td>
<td>URL to the map</td>
<td>A personalized assistant agent is created and associated with the user. The system does the proximity calculation</td>
<td></td>
</tr>
<tr>
<td>Route</td>
<td>Breakdown and service provider Location</td>
<td>Driving direction</td>
<td>A personalized assistant agent is created and associated with the user</td>
<td>User selects the service provider</td>
</tr>
</tbody>
</table>
Design: Acquaintance Model

- **Acquaintance models** define the communication links that exist between agent types. They do not define what messages are sent or when messages are sent. They indicate that communication pathways exist.

- The purpose of acquaintance model is to identify any potential communication bottlenecks, which may cause problems at run-time. ← helpful when testing the system

- An agent acquaintance model is a graph, with nodes in the graph corresponding to agent types and arcs in the graph corresponding to communication pathways.

- Agent acquaintance models are directed graphs: \( A \rightarrow B \) indicates that **A** will send messages to **B**, but not necessarily that **B** will send messages to **A**.
Example: Acquaintance Model

![Diagram showing connections between Geocode, RenderMap, Route, and PersonalAssistant]

- Geocode
- RenderMap
- Route
- PersonalAssistant

far@ucalgary.ca
Design Process

1. Create an agent model:
   - Aggregate roles into agent types, and refine to form an agent type hierarchy;
   - Document the instances of each agent type using instance annotations.

2. Develop a services model, by examining protocols and safety and liveness properties of roles.

3. Develop an acquaintance model from the interaction model and agent model.
Extensions (Future Works)

- Dealing with truly open systems, in which agents may not share common goals (liveness property). This class of systems represents the most important application area for multi-agent systems, and it is therefore essential that the methodology be able to deal with it. case of competition

- Notion of an organizational structure: at the moment, such structures are only implicitly defined within the methodology via the role, interaction and acquaintance models. However, direct, explicit representations of such structures will be of value. Representing such structures may be the only way of adequately capturing and understanding the organization’s communication and control structures.
Critics

Evaluate the methodology along with the followings:

- Roles vs. tasks: yes
- Emergent behaviour: n/a
- Knowledge cycle: no
- Interactions vs. intra-actions: yes
- Symbol level communication: no (may be)
- Knowledge completeness: no
- Indeterminism: Decision making: no
Evaluation /1

- **Concepts:** abstract and concrete concepts are clearly defined. (Good)

- **Modeling constructs:** uses a clear and consistent notation with a clear and defined syntax and semantics. None of the abstracts and concrete concepts or the models are in conflict and they can be interpreted unambiguously. (Good)

- **Scalability:** Roles are not hierarchical. The role model provides strictly one level of abstraction for the developers to conceptualize the system. The methodology does not facilitate iterative refinement of the system through different levels of abstraction. As a result, GAIA does not scale up well to handle complex systems. (Poor)
Evaluation /2

- GAIA does not have the ability to track dependencies between different models and the underlying code used to implement them.
- GAIA does not directly deal with particular modeling techniques or implementation details.
- GAIA models do not specify what tasks are done in parallel, so it has a weak support for concurrency.
- GAIA offers no mechanisms to model the dynamic reasoning, extension and modification of the social aspects at runtime.
- GAIA assumes complete specification of requirements and does not address the requirement gathering phase.
Evaluation /3

- **Limited process life-cycle coverage:** GAIA is intended to allow an analyst to go systematically from a statement of requirements to a design that is sufficiently detailed that it can be implemented directly, but ignoring implementation, testing/debugging, deployment and maintenance details.

- **Limited process management:** GAIA does not address any management and/or cost estimation aspects involved in software development.
Formal semantics: GAIA does not have a formal semantics. A successful methodology is one that is not only of pragmatic value, but one that also has a well-defined, unambiguous formal semantics. While the typical developer need never even be aware of the existence of such a semantics, it is nevertheless essential to have a precise understanding of what the concepts and terms in a methodology mean.
Conclusions

- GAIA is a “popular” methodology for analysis and design of agent-based systems.
- The key concepts in this methodology are **roles**, which are associated with **responsibilities**, **permissions**, and **protocols**. Roles can interact with one another in certain institutionalised ways, which are defined in the protocols of the respective roles.
want a break?
The Tropos Development Methodology
Tropos

- *Tropos* is derived from the Greek *trope* which means “easily changeable” or “easily adaptable.”
- The Tropos Software Methodology was first presented in 2001 at the 5th IEEE International Symposium on Requirements Engineering where the authors proposed a new specification language, called Formal Tropos, that offers the primitive concepts of early requirements frameworks.
Fundamentals

- Tropos has two main features:
  - Agents, goals, plans, and other knowledge level concepts are fundamental primitives in their notation and used informally during the complete software development process.
  - Requirements analysis and specification plays an important role when the system that has to be developed is analyzed with respect to its intended environment.
Comparison

- Tropos is used during the whole software development process.

<table>
<thead>
<tr>
<th>Early Requirements</th>
<th>Late Requirements</th>
<th>Architectural Design</th>
<th>Detailed Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropos</td>
<td>i*</td>
<td>Kaos</td>
<td>Gaia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAll and Mase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AUML</td>
</tr>
</tbody>
</table>
The phases during software development that are covered by Tropos are the following:

- **Early Requirements.** This phase includes identification of relevant stakeholders (actors) with respective objectives (goals).

- **Late Requirements.** The target system is introduced as another actor and is related to stakeholder actors in terms of actor dependencies (these indicate the obligations of the system towards its environment).

- **Architectural Design.** Additional system actors are introduced and assigned to goals or subtasks of system goals or system tasks.

- **Detailed Design.** All system actors are described in detail (including specifications of communication and coordination); output is a Tropos specification.

- **Implementation.** The Tropos specification is transformed to a skeleton for implementation (in detail, this is done by mapping from Tropos constructs to those constructs of an agent programming platform; later code is added to the skeleton).
Conceptual Entities

- Symbolic representation of the main concepts of *Tropos*
Conceptual Entities /2

- **Actor.** An actor is a generalization of a software agent. This concept models a physical agent or a role (abstract characterization of behavior of an actor within a certain context) or position (set of roles played by one agent).

- **Goal.** This concept is a representation of strategic interests of an actor. *Tropos* makes a difference between goals having no clear-cut definition or criteria (softgoals) and hardgoals.

- **Dependency.** A dependency exists between two actors to indicate that one of the actors (*dependeer*) depends on the other (*dependee*) to achieve its goals (or to execute some plan, or deliver a resource). The set of a goal, plan, and resource is called *dependum*.

- **Resource.** Physical or informational entities that one actor wants and another can deliver are represented by resources.

- **Plan.** The concept of a plan represents a way of satisfying a goal.

- **Capability.** This means the capability of an actor to define, choose and execute a plan to fulfill a goal in a given operating environment.

- **Belief.** A belief represents the actor’s knowledge of the world.
Tropos Models

- Several models can be built using Tropos conceptual entities
1. Actor & Dependency Model

- These types of models result from the analysis of social and system actors and their goals and dependencies for the achievement of goals.

- The models are built in the early requirements phase when application domain stakeholders, their intentions and their inter-dependencies are identified.

- Actor models may be extended during the late requirements phase by adding the future system itself as another actor, along with its inter-dependencies with social actors.

- In architectural design level actor and dependency models provide a more detailed structure of the future system actor and its internal structure.
Dependency between two actors: one actor depends on the other to attain some goal, execute some plan or deliver a resource. First actor is called depender, second one is dependee. The object around which the dependency centers is called dependum.
2. Goal and Plan Models

- A designer can analyze goals and plans from an actor’s perspective using three techniques:

  - **Means-ends analysis.** Refining a goal into subgoals to identify plans, resources, and softgoals (which provide means for achieving the goal at the end).

  - **AND/OR (de)composition.** Allows a combination of AND and OR (de)compositions of a root goal into subgoals (creates a goal structure).

  - **Contribution analysis.** To point out goals contributing negatively or positively in reaching the root goal being analyzed. It is a special case of means-end analysis where means are always goals.
Goal and Plan Models

- Samples of *Means-End Analysis*, *OR-Decomposition*, *And-Decomposition*, *Contribution*, and *Actor perspective*.

![Diagram of Goal and Plan Models](image-url)
Goal and Plan Models

- Sample goal diagram for the actor MediaShop
Tropos: Process

- *Tropos* is a generic design process which focuses on a goal analysis for different actors. The process is a non-deterministic concurrent algorithm (including a completeness criterion). During the analysis process the design flows from external to system actors through a goal analysis and delegation.

- The process starts with actors with a list of related goals (and softgoals) and the root goals are analyzed from the actor’s perspective. Generated subgoals are delegated to other actors or the actor itself when it deals with the goals itself. This analysis is carried out concurrently with respect to each root goal. The process may require the introduction of new actors which are delegated to goals and/or tasks. The process ends when all goals have been dealt with to the satisfaction of actors.
Formal Description of Process

global actorList, goalList, agenda, dependencyList, capabilityList, goalGraph,
procedure rootGoalAnalysis(actorList, goalList, goalGraph)
begin
  rootGoalList = nil;
  for actor in actorList do
    for rootGoal in goalList(actor) do
      rootGoalList = add(rootGoal, rootGoalList);
      rootGoal.actor = actor;
    end;
  end;
end;
concurrent for
  rootGoal in rootGoalList do
    goalAnalysis(rootGoal, actorList)
end concurrent for;
if not[satisfied(rootGoalList, goalGraph)] then fail;
end procedure
**Formal Process: Explanation**

<table>
<thead>
<tr>
<th><strong>actorList</strong></th>
<th>Finite set of actors.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>goalList</strong></td>
<td>List of goals for actor (in goalList(actor)).</td>
</tr>
<tr>
<td><strong>agenda</strong></td>
<td>List of goals actor has undertaken to achieve personally (initially empty).</td>
</tr>
<tr>
<td><strong>dependencyList</strong></td>
<td>List of dependencies between actors.</td>
</tr>
<tr>
<td><strong>capabilityList</strong></td>
<td>Includes &lt;goal,plan&gt; pairs indicating means by which the actor can achieve goals.</td>
</tr>
<tr>
<td><strong>goalGraph</strong></td>
<td>Stores representation of goal graph that has been generated so far by the design process (initially contains all root goals of all initial actors with no links among them).</td>
</tr>
<tr>
<td><strong>rootGoalAnalysis()</strong></td>
<td>Conducts concurrent goal analysis for every root goal.</td>
</tr>
<tr>
<td><strong>satisfied()</strong></td>
<td>Checks whether all root goals in goalGraph are satisfied.</td>
</tr>
<tr>
<td><strong>goalAnalysis()</strong></td>
<td>Conducts concurrent goal analysis for every sub goal of a given root goal.</td>
</tr>
</tbody>
</table>
Meta-metamodel level. The *meta-metamodel level* is the basis for the extensions of the metamodel. It contains the language primitives that allow the inclusions of constructs (basic language, structural elements, e.g. attributes, entities).

Metamodel level. The *metamodel level* provides constructs for modeling knowledge level entities and concepts (knowledge level notations, e.g. actor, goal, dependency).

Domain model level. The *domain model level* provides a representation of entities and concepts of specific application domains which are built as instances of the metamodel constructs (application domain entities).

Instance Model level. The *instance model level* contains instances of the domain model.
## Tropos: Modeling Language

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meta-Metamodel</td>
<td>Specifies language structural elements</td>
<td>Attribute, Entity</td>
</tr>
<tr>
<td>Metamodel</td>
<td>An instance of the meta-metamodel. Defines knowledge level notions</td>
<td>Actor, Goal, Plan</td>
</tr>
<tr>
<td>Domain</td>
<td>An instance of the metamodel. Models application domain entities</td>
<td></td>
</tr>
<tr>
<td>Instance</td>
<td>Instantiates domain model elements</td>
<td></td>
</tr>
</tbody>
</table>
At the design level several design patterns, such as *Pair* pattern and *Mediation* patterns can be used to implement the system.

- The *Pair* patterns – such as booking, call-for-proposal, subscription, or bidding – describe direct interactions between negotiating agents.

- The *Mediation* patterns – such as monitor, broker, matchmaker, mediator, embassy, or wrapper – feature intermediary agents that help other agents to reach an agreement on an exchange of services.

- Etc.
Conclusions

- Tropos is one of the “better” methodologies in terms of blending agent-based technology and contemporary software engineering practice.
The MASSIVE Development Method for Multiagent Systems

Jurgen Lind
Abstract

- MASSIVE is a pragmatic process model for the development of multiagent systems based on the combination of standard software engineering techniques.
- The development method features a product model to describe the target system, a process model to construct the product model and an institutional framework that supports learning and reuse over project boundaries.
- The approach has evolved over several years and it has been (successfully?) applied and refined in various industrial research projects.
Introduction

- Late 90’s: The strong need for agents and agent systems development methods in an industrial context is exemplified by the major European telecommunication companies that have launched a joint research project to foster the definition of a development method for agent applications.
Software Development Methods

- A software development method encompasses a notation, whose purpose is to provide a common means of **what** to be developed and specifying **how** (and **when**) certain artifacts should be produced.

- Recently, several development methods for multiagent system have been proposed. However, most of these methods feature a *biased* and *technology-driven*, *model-oriented* and *sequential approach* towards the development of multiagent applications that is not always adequate for this class of systems.
Requirements-driven vs. Technology-driven

- Most of the related development methods assume that the technology, i.e. a multiagent approach, for the target system has already been set and that the system must be designed and implemented according to this technology.

- In a requirements-driven development approach, the requirements of the target system determine the technology that is to be used for the system design and implementation and not vice-versa.

- Thus, a requirements-driven approach starts earlier in the life cycle of the target system when the fundamental characteristics of the problem under consideration are analyzed.
**View-oriented vs. Model-oriented**

- The model-oriented approaches develop several distinct models of the target system that must be integrated in order to construct the final system. This integration process is not trivial as the links between the models are usually neither explicitly defined nor obvious.

- A view-oriented method deals with the system as a whole and uses different perspectives on the entire system as fundamental abstraction. The system is not decomposed but rather viewed from different angles with different foci on particular aspects.

- The advantage of this approach is that it avoids the integration process because the model is always consistent from any view as changes in one view are always directly propagated to other views.
Iterative vs. Sequential

- A sequential process model has the assumption, that everything will work out as desired and that all documents are complete, correct and consistent once they are created. However, this assumption is unlikely that it holds in the case of multiagent systems.

- A process model that is to be used for multiagent applications should respect the fact that the system requirements and design documents are incomplete.

- It is a good idea to iterate the processes that were applied to construct the documents and to evaluate the results of each iteration according the quality measures defined early in the development process.

- This approach reduces the risk of errors because of the repeated quality evaluations.
Balanced vs. Biased

- Traditional SE methods are often biased towards the model aspects of the problem in question and neglect the code part of the resulting system.

- Novel trends in the SE community and especially the Object-Oriented feature a point of view that treats the model and the code that implements the model as equally important parts of the whole.

- In MASSIVE this point-of-view is adopted and a development method that clearly relates the parts of the model and the matching parts of the code is presented.
The MASSIVE framework provides a notation in the form of one or more

- **Product models** that are used to describe the design and/or the implementation of a particular system.
- **Process models** that guide the designer in constructing the product models according to the user requirements.
- **Institutional framework** that tells the project management how the design process must be configured to yield the best possible result.
MASSIVE Views

- A view is a perspective of the model that is meaningful to specific stakeholders.
- A view allows stakeholders to see the system from a perspective most meaningful to them.
- The MASSIVE method is based on *seven views* where each view represents a set of conceptually linked features.
Comparison: UML Views
A view system for the development of multiagent applications should comply to the following requirements:

- Separation: cohesion and coupling; decomposability
- Coverage: integrity
- Flexibility
- Size: Granularity
- Naming
1. Task View

- In the Task view, the functional aspects of the target system are analyzed and a task hierarchy is generated that is then used to determine the basic problem solving capabilities of the entities in the final system.

- Furthermore, the non-functional requirements (e.g., performance, compatibility, security, reliability, usability) of the target system are defined and quantified as far as possible.

- **This view does not assume that a multiagent approach is used for the final system.**

- **Example:** In a compiler application, the functional requirement is that the system translates a program specified in a high-level language to a particular assembly language. The quality of the resulting code or the maximal tolerable time for the compilation are non-functional requirements and the basic problem solving capabilities are lexical analysis or code generation.
2. Environment View

- In environment view, the environment of the target system is analyzed from the *developers perspective* and from the *systems perspective*.

- These two points-of-view usually differ greatly as the developer usually has global knowledge whereas the system has only local knowledge.

- **Example:** In a multiagent system, an individual agent has access to only parts of the environment and resources and there are ongoing activities that cannot be perceived by the agent.
3. Role View

- Role view determines the functional aggregation of the basic problem solving capabilities according to the physical constraints of the target system.

- A role is an abstraction that links the domain dependent part of the application with the agent technology that is used to solve the problem under consideration.

- An agent consists of:
  - One or more role descriptions and
  - An architecture that is capable of executing these role models
Interaction view is the fundamental concept within a system that is comprised of multiple independent entities that must coordinate themselves in order to achieve their individual as well as their joint goals.

In this view, interaction within the target system is seen as a generalized form of conflict resolution that is not limited to a particular form such as communication.

Several generic forms of interaction that can be instantiated in a wide variety of contexts are discussed and the developer is encouraged to analyze the target problem with respect to the applicability of these generic forms before designing new forms.
5. Society View

- A society is a structured collection of entities that pursue a common goal.
- The goal of this view is to classify the society that either pre-exists within the organizational context of the system or that is desirable from the point of view of the system developer.
- A society model is developed that is consistent with the roles within the society and that achieves the defined goals.
- The final structure of the agent society (flat or clustered) thus depends on the precedence of the quality measures (quality of the solution vs. efficiency) that is taken by the system designer.
6. Architectural View

- The architectural view is a projection of the target system onto the fundamental structural attributes with respect to the system design. The major aspects that are dealt with in this view are the system architecture as a whole and the agent architecture.

- **System architecture** is described according to various considerations and includes things such as the agent management or database integration.

- **Agent architecture** is characterized according to the requirements of the problem to be solved.

- System developers should at first try to select one of the numerous existing architectures (platforms) before trying to develop a new architecture from scratch.
7. System View

- System view deals with aspects that affect several of the other views at the same time or even the system as a whole.

- System view handles: user interface that controls the interaction between the system and the user(s) on a task specific as well as the task independent level.

- Task specific aspects are usually input specification and output presentation.

- Task independent aspects deal with the visualization of the system activities in order to enable the user to follow the ongoing computations and interactions.

- Other aspects that are treated in this view are: system-wide error-handling strategy; performance engineering; and system deployment once it has been developed.
Iterative View Engg /1

- Iterative view engineering is a product centered SE process model that combines *Round-trip Engineering* and *Iterative Enhancement*.
- Iterative Enhancement is a process model that is similar to the Waterfall Model except that the Waterfall Model needs a complete and stable requirements definition.
- Iterative Enhancement model operates in several cycles on a partitioned and incomplete system model. The incomplete model is iteratively enhanced in each cycle.
Iterative View Engg /2

- See the Iterative Enhancement as the macro process that encapsulates several micro processes that are used to construct individual parts of the model.

- In the example figure, product model consists of three parts and each part is iteratively constructed with another process model. The vertical boxes represent a snapshot of the entire model at particular points in time $t_1$ and $t_2$ where the parts of the model can be in different process stages.
Iterative View Engg  /3
The Iterative view engineering process model is the following:

- Initially, the model and the implementation are empty. In a first cycle of iterative enhancement (1), the designer specifies the first version of the model probably by using different micro-models for each view.

- In step (2), parts of the initial model are implemented and if an error occurs during the construction phase, the model has to be refined (3) until it can be implemented.

- Next, the initial implementation is tested. During the test phase, it often turns out that the model specification was incomplete and the implementation must be changed or enhanced (4) in order to meet the intentions of the initial model.
After the test and enhancement phase of the implementation is complete, the results must be integrated into the model in an analysis step (5).

If the implementation cannot be re-engineered (e.g. because the changes of the code are incompatible with some basic requirements of the model) the implementation must be changed (6) in order to comply to the modeling language.

After this step, the next cycle is executed until the entire system is fully implemented.
The Experience Factory (BASILI, 1994) is a conceptual framework that supports systematic learning within an organization.

The basic idea underlying the Experience Factory is closely related to case-based reasoning and operates on similarities of software development processes.

The major advantage of the Experience Factory is that it provides a single interface to the knowledge that exists within an organization and a central authority that collects and updates this knowledge.
Experience Factory /2
MASSIVE Development Model

Experience Factory
The core of the MASSIVE method is the view system.

The main idea of the process model is an iterative view engineering approach that is itself based on iterative enhancement and round-trip engineering.

These micro-models as well as the overall process model are in no way fixed for the entire lifetime of the project model, they are also subject to changes and refinements during the course of time.

In order to preserve these adaptations and to make them accessible to others, the process model and the product model are both embedded into a larger organizational structure called the Experience Factory.

The Experience Factory provides the formal framework for a permanent learning process that takes place over project boundaries and that eventually models the multiagent experience of organization in terms of specific product and process models for various domains.
MASSIVE Model in UML
Conclusions

- The author presented a pragmatic software engineering development (process) model that is supposed to be specialized for the development of multiagent systems.
- The model features a view oriented product model, an iterative process model and an institutional framework that supports organizational learning and reuse over project boundaries.
Critics and Questions

- Is the view system adequate?
- Assigning roles and responsibilities
- Scaling issues
- Performance and testing issues
- Measurement issues
- Decentralized knowledge bases
- Uncertainty management issues
Critics

Evaluate the methodology along with the followings:

- Roles vs. tasks: yes/no
- Emergent behaviour: no
- Knowledge cycle: no
- Interactions vs. intra-actions: yes
- Symbol level communication: no
- Knowledge completeness: yes
- Indeterminism: Decision making: no
A Statistical Approach for Comparing Agent Based Methodologies
Motivation

- Increased interest in MAS → large number of AOSE methodologies
- AOSE methodologies all utilize agent concepts and offer techniques
  - Suitable for different stages of software development process
  - Tailored to different levels of abstraction
Motivation /2

- From literature, we enumerated more than 30 AOSE methodologies
- Some methodologies do not have adequate capabilities to model the properties of all types of agents
- Some methodologies are developed for special purposes
- Some methodologies do not have sufficient documentation or presented in languages other than English
Research Question

- What is the best AOSE methodology for a given set of requirements?

- The ambiguity and lack of standardization associated with the existing AOSE methodologies makes it difficult to take decisions to select the most appropriate methodology that best fits a particular agent-oriented application.
Related Works /1

- In the literature, a limited number of studies have been carried out to challenge the problem of evaluating and comparing AOSE methodologies through a number of frameworks.

- However, they may not necessarily follow a reliable approach assuring effective evaluation and accurate decisions.
Most of the evaluation studies pay attention to highlight the best features of the presented methodologies

Most of the evaluation studies that we found have been conducted by the developers of methodologies. This may create a bias that makes developers unable to recognize drawbacks with their own developed methodologies.
The main objective of this research is to devise a technique for \textit{evaluating AOSE methodologies} and come up with a set of facts that support the process of \textit{selecting the most appropriate methodology} to accommodate the anticipated criteria of new agent-based applications.
Research Questions

1. What are the major criteria upon which we can describe and evaluate an AOSE methodology?

2. What are the different features or attributes that specify each of these criteria?

3. How can we quantify and assess the evaluation criteria of AOSE methodologies, as well as their descended attributes?
Research Questions

4. What are the software metrics that can be used to evaluate an AOSE methodology?

5. How can we determine whether significant differences exist between the evaluated AOSE methodologies?

6. How can we rank the evaluated AOSE methodologies in order to select the most appropriate methodology for a given MAS application?
Methodology  /1

- Our methodological approach can be divided into the following 4 major steps:

**Step 1: Surveying Available Methodologies**

- We started the research by conducting a comparative survey to review most of the available AOSE methodologies that are presented in literature. As a result, we came up with 31 methodologies that have reasonable documentation to study.
Step 2: Selecting Qualified Methodologies

- We defined a qualified methodology as a methodology that meets the following prerequisites:
  - The methodology has reasonable documentation
  - The methodology is fairly known to the agent community
  - The methodology is not a special-purpose one rather, it has a reasonable domain of applicability
Methodology

By reviewing the available 31 methodologies against these assumptions, the following 9 methodologies have been selected as the most qualified ones to carry on through the evaluation process:

- **Gaia** [Wooldridge 2000], **MaSE** [DeLoach 2002], **Tropos** [Giunchiglia 2002], **Agent-SE** [Far 2002], **MASSIVE** [Lind 2002], **Prometheus** [Padgham 2002, Dam, 2003], **MESSAGE** [Evans 2001], **MAS-CommonKADS** [Iglesias 1997], and **PASSI** [Cossentino 2005]
Step 3: Applying MWAF

- We proposed the Multidimensional Weighted-Attributes Framework (MWAF), as a general-purpose framework that utilizes statistical methods for evaluating software in general, and AOSE methodologies in particular.

Step 4: Statistical Analysis
Methodology: MWAF

Components of MWAF:
1. Dimensions
2. Attributes
3. Parameters
Rates & Weights

- “Rate” reflects the extent the attribute is supported in the methodology. (how the attribute is effectively supported?)

- “Weight” reflects the extent that the attribute is important to agent-based software design and development. (how important is to MAS development?)
Methodology: MWAF /2

The main idea of MWAF is to define the most common and important criteria (we referred to as dimensions) of the software being evaluated,

- Identifying the attributes that describe each of the dimensions, and then

- Evaluating each dimension through its attributes against all the potential AOSE methodologies that are given for evaluation.
We identified 8 dimensions that best describe an AOSE methodology.

We can break down these dimensions into the hierarchical level of their descended attributes, as follows:
Methodology: MWAF /4

- Dimension 1: Agency-relates
Methodology: MWAF

- Dimension 2: Modeling

  - 2.1 Notation
  - 2.2 Ease of Use and Understanding
  - 2.3 Expressiveness
  - 2.4 Abstraction
  - 2.5 Consistency
  - 2.6 Concurrency
  - 2.7 Traceability
  - 2.8 Derivation and Reusability
Methodology: MWAF /6

- Dimension 3: Communication

  3.1 Local Communication (Basic Sociability)
  - 3.1.1 Cooperation
  - 3.1.2 Coordination
  - 3.1.3 Competition
  - 3.1.4 Negotiation

  3.2 Wide Communication (Advanced Sociability)
  - 3.2.1 Interaction with the External Environment
    - 3.2.1.1 Agent-based User Interface
    - 3.2.3 Subsystems Interaction
    - 3.2.4 Bio-induction
Methodology: MWAF

- **Dimension 4:**
  - Process

  - 4.1 Development Lifecycle
  - 4.2 Refinability
  - 4.3 Managing Complexity
  - 4.4 Ease of Use and Understanding
  - 4.1.1 Architectural Design
  - 4.1.2 Detailed Design
  - 4.1.3 Verification and Validation
Methodology: MWAF

- **Dimension 5:**
  - **Upgrade**

  - 5.1 Modifiability
  - 5.2 Scalability
  - 5.3 Open Systems Support
  - 5.4 Adaptability – Dynamic Structure
  - 5.5 Integrateability
Methodology: MWAF

- Dimension 6:
  Application

- 6.1 Applicability
- 6.2 Maturity
- 6.3 Field history
- 6.4 Cost concerns
Methodology: MWAF /10

- **Dimension 7:** Supporting Properties

  - 7.1 Ontology
  - 7.2 Security
  - 7.3 Collaborative Services
**Methodology: MWAF /11**

- **Dimension 8:**
  - **User Perception**

  - **8.1 Perceived Ease of Use**
  - **8.2 Perceived Usefulness**
  - **8.3 Intention to Use**

far@ucalgary.ca 140
## Methodology: MWAF /12

<table>
<thead>
<tr>
<th>d/n</th>
<th>a/n</th>
<th>Sub-attributes</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
<th>P11</th>
<th>P12</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>A1</td>
<td>1.1.1 Organization</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>8.417</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>1.1.2 Mobility</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>7.667</td>
</tr>
<tr>
<td></td>
<td>A3</td>
<td>1.2.1 Autonomy</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>9.083</td>
</tr>
<tr>
<td></td>
<td>A4</td>
<td>1.2.2 Reactivity</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7.250</td>
</tr>
<tr>
<td></td>
<td>A5</td>
<td>1.2.3 Reasoning</td>
<td>9</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8.167</td>
</tr>
<tr>
<td></td>
<td>A6</td>
<td>1.2.4 Temporal continuity</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>8.083</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A7</td>
<td>1.3.1 Beliefs</td>
<td>6</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>8.333</td>
</tr>
<tr>
<td></td>
<td>A8</td>
<td>1.3.2 Desires</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>7.250</td>
</tr>
<tr>
<td></td>
<td>A9</td>
<td>1.3.3 Intentions</td>
<td>5</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>7.083</td>
</tr>
<tr>
<td></td>
<td>A10</td>
<td>1.4 Learning ability</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>9.000</td>
</tr>
<tr>
<td>D2</td>
<td>A11</td>
<td>2.1 Notation</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>8.250</td>
</tr>
<tr>
<td></td>
<td>A12</td>
<td>2.2 Ease of use</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>0</td>
<td>5.170</td>
</tr>
<tr>
<td></td>
<td>A13</td>
<td>2.3 Expressiveness</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>7.917</td>
</tr>
<tr>
<td></td>
<td>A14</td>
<td>2.4 Abstraction</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7.500</td>
</tr>
<tr>
<td></td>
<td>A15</td>
<td>2.5 Consistency</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>8.750</td>
</tr>
<tr>
<td></td>
<td>A16</td>
<td>2.6 Abstraction</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7.083</td>
</tr>
<tr>
<td></td>
<td>A17</td>
<td>2.7 Traceability</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>6</td>
<td>9</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>8</td>
<td>7</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A18</td>
<td>2.8 Derivation and reuse</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8.083</td>
</tr>
<tr>
<td>D3</td>
<td>A19</td>
<td>3.1.1 Cooperation</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>7.333</td>
</tr>
<tr>
<td></td>
<td>A20</td>
<td>3.1.2 Coordination</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>7</td>
<td>7.917</td>
</tr>
<tr>
<td></td>
<td>A21</td>
<td>3.1.3 Competition</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>7.667</td>
</tr>
<tr>
<td></td>
<td>A22</td>
<td>3.1.4 Negotiation</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>7.417</td>
</tr>
<tr>
<td></td>
<td>A23</td>
<td>3.2.1 External Interaction</td>
<td>10</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>8.333</td>
</tr>
<tr>
<td></td>
<td>A24</td>
<td>3.2.2 Agent-based user int.</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>7.583</td>
</tr>
<tr>
<td></td>
<td>A25</td>
<td>3.2.3 Subsystems inter.</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7.000</td>
</tr>
<tr>
<td></td>
<td>A26</td>
<td>3.2.4 Bio-induction</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>10</td>
<td>4.420</td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>A27</td>
<td>4.1.1 Architectural design</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7.167</td>
</tr>
<tr>
<td></td>
<td>A28</td>
<td>4.1.2 Detailed design</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>8</td>
<td>8</td>
<td>/</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>10</td>
<td>1.911</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A29</td>
<td>4.1.3 Verification/validat.</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>8.083</td>
<td></td>
</tr>
</tbody>
</table>
**Methodology: MWAF**

| A30  | 4.2 Refinability | 8  | 8  | 7  | 6  | 9  | 8  | 7  | 6  | 6  | 7  | 8  | 8  | 7.333 |
| A31  | 4.3 Managing complexity | 9  | 7  | 9  | 7  | 9  | 10 | 10 | 7  | 10 | 9  | 8  | 8  | 8.583 |
| A32  | 4.4 Ease of use | 0  | 0  | 0  | 0  | 9  | 10 | 9  | 8  | 9  | 8  | 9  | 0  | 5.170 |
| D5   | A33  | 5.1 Modifiability | 8  | 6  | 0  | 5  | 1  | 0  | 8  | 8  | 9  | 8  | 8  | 0  | 5.080 |
| A34  | 5.2 Scalability | 9  | 6  | 0  | 5  | 1  | 0  | 0  | 8  | 5  | 6  | 7  | 0  | 3.920 |
| A35  | 5.3 Open system support | 8  | 9  | 0  | 9  | 8  | 0  | 0  | 8  | 5  | 6  | 6  | 0  | 4.920 |
| A36  | 5.4 Adapt-ability/dynamic | 7  | 9  | 1  | 7  | 8  | 0  | 0  | 7  | 5  | 7  | 7  | 0  | 4.830 |
| A37  | 5.5 Integrateability | 6  | 9  | 0  | 7  | 8  | 0  | 0  | 8  | 0  | 6  | 8  | 6  | 4.830 |
| D6   | A38  | 6.1 Applicability | 9  | 9  | 8  | 8  | 7  | 8  | 7  | 6  | 9  | 6  | 7  | 10 | 7.833 |
| A39  | 6.2 Maturity | 7  | 7  | 7  | 5  | 6  | 8  | 9  | 8  | 10 | 8  | 5  | 9  | 7.417 |
| A40  | 6.3 Field history | 9  | 9  | 7  | 5  | 9  | 9  | 9  | 6  | 8  | 7  | 7  | 7  | 7.667 |
| A41  | 6.4 Cost concerns | 6  | 7  | 7  | 6  | 8  | 7  | 7  | 7  | 9  | 7  | 9  | 7.250 |
| D7   | A42  | 7.1 Ontology | 5  | 9  | 0  | 7  | 7  | 9  | 8  | 0  | 9  | 8  | 0  | 6  | 5.670 |
| A43  | 7.2 Security | 5  | 8  | 0  | 6  | 9  | 9  | 5  | 0  | 7  | 7  | 0  | 7  | 5.250 |
| A44  | 7.3 Collaborative services | 5  | 6  | 0  | 7  | 8  | 8  | 7  | 0  | 9  | 7  | 0  | 0  | 4.750 |
| D8   | A45  | 8.1 Perceived ease of use | 9  | 9  | 9  | 6  | 7  | 7  | 7  | 6  | 8  | 9  | 8  | 7  | 7.667 |
| A46  | 8.2 Perceived usefulness | 9  | 9  | 9  | 7  | 9  | 9  | 9  | 8  | 8  | 9  | 9  | 9  | 8.667 |
| A47  | 8.3 Intention to use | 6  | 9  | 5  | 7  | 8  | 10 | 7  | 8  | 9  | 9  | 9  | 10 | 8.083 |

**

- Scores less than 5 are in bold.
- Average weights less than 7.0 are in italic bold.
- d/n refers to dimension number
- a/n refers to attribute number
Statistical Analysis /1

Step 1. Selecting Participants

- There were 12 participants in this study, all of whom had adequate knowledge and experience in software development and MAS (min 4 years software development experience; passed AOSE course; designed and developed at least one MAS)

- The participants were provided sufficient documentation about the methodologies, clear instructions about the experiment, and equal amount of time to complete their task using well-prepared surveying questionnaire
Statistical Analysis /2

Step 2. Selecting the appropriate statistical model

- **Complete Random Design (CRD)**
  - Requires 36 participants to evaluate the 9 methodologies; each methodology will be evaluated by 4 subjects.

- **Randomized Complete Block Design (RCBD)**
  - Each evaluator should assess a complete block, i.e. 9 methodologies, and the design will probably be more effective because we will have more replications, $12 \times 9 = 108$ instead of 36 (i.e. 9 replica/treatment).
  - Hard to implement; not all the participants are familiar with the whole set of methodologies.

** We decided to have at least 4 replica **
The statistical approach we adopted in MWAF depends on applying the Analysis of Variance (ANOVA) procedure to the Balanced Incomplete Block Design (BIBD) model to test for the significant differences in the mean effectiveness of each individual dimension among the evaluated methodologies.
Anova

- The basic idea of ANOVA is to compare the variability of the observations between groups to the variability within groups, such that:
  - if the variability between groups is smaller than the variability within groups, this contributes to supporting the null hypothesis that the means of the different groups are not significantly different; but
  - if the variability between groups is larger than the variability within groups, this discredits the null hypothesis and supports the alternative hypothesis.

- ANOVA is usually used when having 3 or more groups
Balanced Incomplete Block Design (BIBD)

the 12 participants have been used to evaluate the 9 methodologies, such that each participant will evaluate 3 methodologies only

observed rate of effectiveness, \( Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \)

\( \mu \) = overall mean of experiments
\( \alpha_i \) = treatment main effects,
  the deviation from the mean caused by the ith treatment;
  \( \Sigma \alpha_i = 0 \) (fixed effects)
\( \beta_j \) = main effects of the blocking variable,
  the deviation from the mean caused by the jth block;
  \( \Sigma \beta_j = 0 \) (fixed effects)
\( \epsilon_{ij} \) = independent random error \( \sim N(0, \sigma^2) \)
Statistical Analysis: BIBD /2

- **Incomplete:** each participant will not evaluate a complete set of methodologies
- **Balanced:** each block will have the same number of methodologies
- **Additive:** no interactions are considered between blocks and treatments
- **Fixed:** we were limited to narrow down the selection of the qualified methodologies upon our own interest, and not at random from a large number of methodologies
## Treatment Assignment

<table>
<thead>
<tr>
<th>Block (Participants), j</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
<th>M9</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P10</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P11</td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P12</td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

M = Methodology; P = Participant

Treatments (AOSE Methodologies), i
## Testing Model Aptness

<table>
<thead>
<tr>
<th>Test #</th>
<th>Test Type</th>
<th>Instrument(s) Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Outliers</td>
<td>a. Normal probability plot of residuals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Individual value plot of residuals versus independent variable</td>
</tr>
<tr>
<td>2</td>
<td>Normality of residuals</td>
<td>Normal probability plot of residuals</td>
</tr>
<tr>
<td>3</td>
<td>Homogeneity of error variances</td>
<td>a. Residual plots against fitted values</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Bartlett’s test</td>
</tr>
</tbody>
</table>
There was no indication of non-normality, nor was there any evidence pointing to possible outliers; all the residuals fall within the boundaries of the 95% confidence interval.
Outliers & Residuals

- Plot of residuals against the independent variable: Methodology

Individual Value Plot of Standardized Residual vs Methodology

95% CI for the Mean

Standardized Residual vs Methodology

far@ucalgary.ca
Residuals seem to be unstructured and unrelated to the predicted values of the response variable, that is no evidence of heterogeneity among the error variances.

**Residuals Versus the Fitted Values**
(response is Rate)
Model Conclusion

- We repeated these aptness tests for all 6 dimensions.
- All of the plots for all dimensions suggest that ANOVA model is appropriate for the available data.

- What is next? (two step evaluation)
  - Evaluate each dimension
  - Pairwise comparison
Statistical Hypotheses

- **Null hypothesis H₀**
  - There is no significant difference in the mean effectiveness ($τ_i$) of the examined dimension among the evaluated AOSE methodologies, i.e., all the treatment effects are the same.

- **Alternative hypothesis Hₐ**
  - There is a significant difference in the mean effectiveness ($τ_i$) of the examined dimension among the evaluated AOSE methodologies.
The null hypotheses test will allow to check whether there is a significant difference in the effectiveness of a specific dimension among the evaluated methodologies.

If no variability observed, no need to continue the experiment.

However, if the test result is significant, we will go further to carry out multiple pairwise comparisons using Tukey’s HSD (honestly significant diffs!) method to identify which pairs of methodologies are significantly different from the others.
First conclusion:

- The criteria represented by Dimensions 1-6, as evaluated by their rate of effectiveness, has shown significant differences among the 9 evaluated AOSE methodologies.

- Therefore pairwise comparison is conducted for all dimensions.
## Evaluation Results

### Dimension 1: Agency

<table>
<thead>
<tr>
<th></th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
<th>M9</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>M2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Dimension 2: Modeling

<table>
<thead>
<tr>
<th></th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>M5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Legend

- M1: Gaia
- M2: MaSE
- M3: Tropos
- M4: Agent-SE
- M5: MASSIVE
- M6: Prometheus
- M7: MESSAGE
- M8: MARS-CommonKADS
- M9: PASSI

1 = significant difference  
0 = insignificant difference
### Dimension 3: Communications

<table>
<thead>
<tr>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Dimension 4: Process

<table>
<thead>
<tr>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation Results /3

<table>
<thead>
<tr>
<th>Dimension 5: Application</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension 6: User Perception</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>M3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Indicates that our subjects could not reveal significant differences among the methodologies for this dimension!
Second conclusion: Methodologies ranked according to estimated mean of effectiveness of each dimension

<table>
<thead>
<tr>
<th>Rank</th>
<th>Methodology</th>
<th>Est. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td>M1: Gaia</td>
<td>$\hat{\mu}_1 = 6.494$</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>M2: MaSE</td>
<td>$\hat{\mu}_2 = 5.861$</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>M3: Tropos</td>
<td>$\hat{\mu}_3 = 5.835$</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>M9: PASSI</td>
<td>$\hat{\mu}_4 = 5.713$</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>M8: MAS-Common</td>
<td>$\hat{\mu}_5 = 5.549$</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>M7: MESSAGE</td>
<td>$\hat{\mu}_6 = 5.524$</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>M4: Agent-SE</td>
<td>$\hat{\mu}_7 = 5.192$</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>M6: Prometheus</td>
<td>$\hat{\mu}_8 = 5.049$</td>
</tr>
<tr>
<td><strong>9</strong></td>
<td>M5: MASSIVE</td>
<td>$\hat{\mu}_9 = 4.684$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>Methodology</th>
<th>Est. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td>M2: MaSE</td>
<td>$\hat{\mu}_1 = 6.593$</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>M9: PASSI</td>
<td>$\hat{\mu}_2 = 6.428$</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>M1: Gaia</td>
<td>$\hat{\mu}_3 = 6.037$</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>M7: MESSAGE</td>
<td>$\hat{\mu}_4 = 5.777$</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>M8: MAS-Common</td>
<td>$\hat{\mu}_5 = 5.560$</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>M5: MASSIVE</td>
<td>$\hat{\mu}_6 = 5.271$</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>M6: Prometheus</td>
<td>$\hat{\mu}_7 = 5.074$</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>M4: Agent-SE</td>
<td>$\hat{\mu}_8 = 4.755$</td>
</tr>
<tr>
<td><strong>9</strong></td>
<td>M3: Tropos</td>
<td>$\hat{\mu}_9 = 4.580$</td>
</tr>
</tbody>
</table>
### Evaluation Results /5

#### Dimension 3: Communication

<table>
<thead>
<tr>
<th></th>
<th>M7: MESSAGE</th>
<th>$\beta_7 = 6.945$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>M8: MAS-Common</td>
<td>$\beta_8 = 6.847$</td>
</tr>
<tr>
<td>3</td>
<td>M9: PASSI</td>
<td>$\beta_9 = 6.560$</td>
</tr>
<tr>
<td>4</td>
<td>M2: MaSE</td>
<td>$\beta_2 = 5.805$</td>
</tr>
<tr>
<td>5</td>
<td>M5: MASSIVE</td>
<td>$\beta_5 = 5.760$</td>
</tr>
<tr>
<td>6</td>
<td>M1: Gaia</td>
<td>$\beta_1 = 5.446$</td>
</tr>
<tr>
<td>7</td>
<td>M3: Tropos</td>
<td>$\beta_3 = 5.099$</td>
</tr>
<tr>
<td>8</td>
<td>M4: Agent-SE</td>
<td>$\beta_4 = 4.619$</td>
</tr>
<tr>
<td>9</td>
<td>M6: Prometheus</td>
<td>$\beta_6 = 4.575$</td>
</tr>
</tbody>
</table>

#### Dimension 4: Process

<table>
<thead>
<tr>
<th></th>
<th>M1: Gaia</th>
<th>$\beta_1 = 6.932$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>M2: MaSE</td>
<td>$\beta_2 = 6.830$</td>
</tr>
<tr>
<td>3</td>
<td>M5: MASSIVE</td>
<td>$\beta_3 = 5.839$</td>
</tr>
<tr>
<td>4</td>
<td>M8: MAS-Common</td>
<td>$\beta_8 = 5.477$</td>
</tr>
<tr>
<td>5</td>
<td>M9: PASSI</td>
<td>$\beta_9 = 5.460$</td>
</tr>
<tr>
<td>6</td>
<td>M6: Prometheus</td>
<td>$\beta_6 = 5.227$</td>
</tr>
<tr>
<td>7</td>
<td>M4: Agent-SE</td>
<td>$\beta_4 = 4.906$</td>
</tr>
<tr>
<td>8</td>
<td>M7: MESSAGE</td>
<td>$\beta_7 = 4.808$</td>
</tr>
<tr>
<td>9</td>
<td>M3: Tropos</td>
<td>$\beta_3 = 4.502$</td>
</tr>
</tbody>
</table>
### Evaluation Results

#### Dimension 5: Application

<table>
<thead>
<tr>
<th>Rank</th>
<th>Tool</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M2: MaSE</td>
<td>6.351</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M1: Gaia</td>
<td>6.301</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M9: PASSI</td>
<td>5.900</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M7: MESSAGE</td>
<td>5.687</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M5: MASSIVE</td>
<td>5.527</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M8: MAS-CommonKADS</td>
<td>5.411</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M3: Tropos</td>
<td>5.398</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M4: Agent-SE</td>
<td>5.368</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M6: Prometheus</td>
<td>5.187</td>
<td></td>
</tr>
</tbody>
</table>

#### Dimension 6: User Perception

<table>
<thead>
<tr>
<th>Rank</th>
<th>Tool</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M2: MaSE</td>
<td></td>
<td>7.340</td>
</tr>
<tr>
<td>2</td>
<td>M1: Gaia</td>
<td></td>
<td>6.672</td>
</tr>
<tr>
<td>3</td>
<td>M7: MESSAGE</td>
<td></td>
<td>6.198</td>
</tr>
<tr>
<td>4</td>
<td>M8: MAS-CommonKADS</td>
<td></td>
<td>6.118</td>
</tr>
<tr>
<td>5</td>
<td>M5: MASSIVE</td>
<td></td>
<td>5.376</td>
</tr>
<tr>
<td>6</td>
<td>M9: PASSI</td>
<td></td>
<td>5.251</td>
</tr>
<tr>
<td>7</td>
<td>M4: Agent-SE</td>
<td></td>
<td>5.193</td>
</tr>
<tr>
<td>8</td>
<td>M3: Tropos</td>
<td></td>
<td>4.928</td>
</tr>
<tr>
<td>9</td>
<td>M6: Prometheus</td>
<td></td>
<td>3.818</td>
</tr>
</tbody>
</table>
It is rather hard to devise an overall ranking for the evaluated methodologies (e.g., weighted averaging, linear optimization, etc.).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/9</td>
<td>M1</td>
<td>M2</td>
<td>M7</td>
<td>M1</td>
<td>M2</td>
<td>M2</td>
</tr>
<tr>
<td>2</td>
<td>8/9</td>
<td>M2</td>
<td>M9</td>
<td>M8</td>
<td>M2</td>
<td>M1</td>
<td>M1</td>
</tr>
<tr>
<td>3</td>
<td>7/9</td>
<td>M3</td>
<td>M1</td>
<td>M9</td>
<td>M5</td>
<td>M9</td>
<td>M7</td>
</tr>
<tr>
<td>4</td>
<td>6/9</td>
<td>M9</td>
<td>M7</td>
<td>M2</td>
<td>M8</td>
<td>M7</td>
<td>M8</td>
</tr>
<tr>
<td>5</td>
<td>5/9</td>
<td>M8</td>
<td>M8</td>
<td>M5</td>
<td>M9</td>
<td>M5</td>
<td>M5</td>
</tr>
<tr>
<td>6</td>
<td>4/9</td>
<td>M7</td>
<td>M5</td>
<td>M1</td>
<td>M6</td>
<td>M8</td>
<td>M9</td>
</tr>
<tr>
<td>7</td>
<td>3/9</td>
<td>M4</td>
<td>M6</td>
<td>M3</td>
<td>M4</td>
<td>M3</td>
<td>M4</td>
</tr>
<tr>
<td>8</td>
<td>2/9</td>
<td>M6</td>
<td>M4</td>
<td>M4</td>
<td>M7</td>
<td>M4</td>
<td>M3</td>
</tr>
<tr>
<td>9</td>
<td>1/9</td>
<td>M5</td>
<td>M3</td>
<td>M6</td>
<td>M3</td>
<td>M6</td>
<td>M6</td>
</tr>
</tbody>
</table>
Simple Overall Rank

Ranking based on the proportional order of each methodology against the evaluated dimensions

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Total weight</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1: Gaia</td>
<td>$\frac{1}{9}[9+7+4+9+9+8]=\frac{46}{9}$</td>
<td>2</td>
</tr>
<tr>
<td>M2: MaSE</td>
<td>$\frac{1}{9}[8+9+6+8+9+9]=\frac{49}{9}$</td>
<td>1</td>
</tr>
<tr>
<td>M3: Tropos</td>
<td>$\frac{1}{9}[7+1+3+1+9+2]=\frac{23}{9}$</td>
<td>6</td>
</tr>
<tr>
<td>M4: Agent-SE</td>
<td>$\frac{1}{9}[3+2+2+3+9+3]=\frac{22}{9}$</td>
<td>7</td>
</tr>
<tr>
<td>M5: MASSIVE</td>
<td>$\frac{1}{9}[1+4+5+7+9+5]=\frac{31}{9}$</td>
<td>5</td>
</tr>
<tr>
<td>M6: Prometheus</td>
<td>$\frac{1}{9}[2+3+1+4+9+1]=\frac{20}{9}$</td>
<td>8</td>
</tr>
<tr>
<td>M7: MESSAGE</td>
<td>$\frac{1}{9}[4+6+9+2+9+7]=\frac{37}{9}$</td>
<td>4</td>
</tr>
<tr>
<td>M8: MAS-CommonKADS</td>
<td>$\frac{1}{9}[5+5+8+6+9+6]=\frac{39}{9}$</td>
<td>3</td>
</tr>
<tr>
<td>M9: PASSI</td>
<td>$\frac{1}{9}[6+8+7+5+9+4]=\frac{39}{9}$</td>
<td>3</td>
</tr>
</tbody>
</table>
Recommended Future Works

1. Unifying AOSE methodologies

- The comparative statistical analysis we carried out in this research can actually be considered as a preliminary step towards unifying AOSE methodologies.

- Attributes receiving higher rate and weight are to be supported in the unified AOSE methodology.
2. Proposing agent-based software metrics

- Later we can evaluate AOSE methodologies to check whether they can deliver standardized scores that can be generalized as agent-based software metrics, or benchmarks.

- This can be useful to evaluate any AOSE methodology by comparing the rates it gives against these metrics. That is, the relative efficiency of the methodology as compared to the AOSE metrics.
Limitations of Research

- Although the evaluation methodology is reliable and the model aptness tests were successful, the results obtained from this research are of "limited generality" mainly due to:
  - Type of experiment (i.e., balanced incomplete block design BIBD vs. complete block design CBD, etc.); and
  - Limited resources of selecting experimental subjects from a relatively small population.
Conclusions

- We exhibited a statistical approach, based on adopting the BIBD model to evaluate the most common AOSE methodologies.
- We identified the most common and important criteria, upon which we evaluate an AOSE methodology.
- By implementing this approach, we come up with a set of significantly tested scores, in terms of population mean responses, that can be determined by applying multiple comparisons, after assuring significant differences among the methodologies.


References


The End!