The SysML/KAOS Domain Modeling Approach

Steve TUENO, Régine LALEAU, Amel MAMMAR, Marc FRAPPIER

1 Introduction

A means of building safe critical systems consists of formally modeling the requirements formulated by stakeholders and ensuring their consistency with respect to application domain properties. This paper proposes a metamodel for an ontology modeling formalism based on OWL and PLIB. This modeling formalism is part of a method for modeling the domain of systems whose requirements are captured through SysML/KAOS. The formal semantics of SysML/KAOS goals are represented using Event-B specifications. Goals provide the set of events, while domain models will provide the structure of the system state of the Event-B specification. Our proposal is illustrated through a case study dealing with a Cycab localization component specification [2]. The case study deals with the specification of a localization software component that uses GPS, Wi-Fi and sensor technologies for the realtime localization of the Cycab vehicle [10], an autonomous ground transportation system designed to be robust and completely independent.

The remainder of this paper is structured as follows: Section 2 briefly describes the SysML/KAOS method. Follows a presentation, in Section 3, of the relevant state of the art on domain modeling in requirements engineering and a comparison of
ontology modeling formalisms. In Section 4, we describe and illustrate our approach to model the domain of a system specified using the SysML/KAOS method.

2 SysML/KAOS

Requirements engineering focuses on defining and handling requirements. These and all related activities, in order to be carried out, require the choice of an adequate means for requirements representation. The KAOS method \cite{10,11}, proposes to represent the requirements in the form of goals, which can be functional or non-functional, through five sub-models of which the two main ones are: the object model which uses the UML class diagram for the representation of domain vocabulary and the goal model for the determination of requirements to be satisfied by the system and of expectations with regard to the environment through a goals hierarchy having strategic goals formulated by stakeholders at the root level. The hierarchy is built through a succession of refinements using different operators: AND, OR and MILESTONE. An AND refinement decomposes a goal into subgoals, and all of them must be achieved to realise the parent goal. Dually, an OR refinement decomposes a goal into subgoals such that the achievement of only one of them is sufficient for the accomplishment of the parent goal. A MILESTONE refinement is a variant of AND refinement which allows the definition of an achievement order between goals. Requirements and expectations correspond to the lowest level goals of the model.

KAOS proposes a structured approach to obtaining the requirements based on expectations formulated by stakeholders. Unfortunately, it offers no mechanism to maintain a strong traceability between those requirements and deliverables associated with system design and implementation, making it difficult to validate them against the needs formulated. The SysML UML profile has been specially designed by the Object Management Group (OMG) for the analysis and specification of complex systems and allows for the capturing of requirements and the maintaining of traceability links between those requirements and design diagrams resulting from the system design phase. Unfortunately, OMG has not defined a formal semantics and an unambiguous syntax for requirements specification. SysML/KAOS \cite{7} therefore proposes to extend the SysML metamodel with a set of concepts allowing to represent requirements in SysML models as KAOS goals.

A functional goal, under SysML/KAOS, describes the expected behaviour of the system once a certain condition holds \cite{11} : \textit{[if CurrentCondition then] sooner-or-later TargetCondition}. SysML/KAOS allows the definition of a functional goal without specifying a \textit{CurrentCondition}. In this case, the expected behaviour can be observed from any system state.

Figure \ref{fig:goal_diagram} is a goal diagram from the Cycab System localization component focused on the purpose of vehicle localization.

To achieve the root goal, which is the localization of the vehicle (LocalizeVehicle), raw localizations must be captured from vehicle sub components (CaptureRawLocalizations) which can be GPS (CaptureGPSLocalization) or Wi-Fi (CaptureWIFILocaliza-
Fig. 1 Excerpt from the localization component goal diagram

3 State of the Art On Domain Modeling in Requirements Engineering

3.1 Existing Approaches

In KAOS [10], the domain of a system is specified by an object model described by UML class diagrams. An object within this model can be an entity if it exists independently of the others and does not influence the state of any other object, an association if it links other objects on which it depends, an agent if it actively influences the system state by acting on other objects or an event if its existence is instantaneous, appearing to impulse an update of the system state. This approach, which is essentially graphic and informal as argued in [12], is difficult to exploit in case of critical systems [13]. Moreover, it does not offer mechanisms for referencing a model within another, which limits the reusability of models.

In [5], author proposes to model the knowledge of the domain through either formulae of first-order logic or ontologies. He considers an ontology as a more structured and extensible representation of domain knowledge.

In [9], the domain model is built around the notions of Concept and Relationship. Each entry in this model consists of an assertion linking two instances of Concept through a Relationship instance. A categorization is proposed for concepts and relationships: a concept can be a function, an object, a constraint, an actor, a
platform, a quality or an ambiguity, while a relationship can be a performative or a symmetry, reflexivity or transitivity relation. However, the proposed metamodel appears to be incomplete. Indeed, it does not allow to represent key elements of domain modeling as for example the cardinality of a relationship or the attributes of a concept. Moreover, it does not allow to establish references between several models, which limits their reuse.

In [13], ontologies are used not only to represent domain knowledge, but also to model and analyze requirements. The proposed methodology is called knowledge-based requirements engineering (KBRE) and is mainly used for detection and processing of inconsistencies, conflicts and redundancies among requirements. In spite of the fact that KBRE proposes to model the domain knowledge through ontologies, the proposal focuses on the representation of requirements and proposes nothing regarding domain modeling. It is in the same vein that the GOORE method is presented in [18].

In [4], authors are interested in a systematic review of the literature related to applications of ontologies in requirements engineering. They end up describing ontologies as a standard form of formal representation of concepts within a domain, as well as of relationships between those concepts. This is equivalent to considering ontologies as a standard for formal modeling of system domain.

These approaches suggest that ontologies are relevant for modeling the domain of a system.

3.2 A Study of Ontology Modeling Formalisms

An ontology can be defined as a formal model representing concepts that can be grouped into categories through generalization/specialization relations, their instances, constraints and properties as well as relations existing between them. Ontology modeling formalisms can be grouped into two categories: Closed World Assumption (CWA) for those considering that any fact that cannot be deduced from what is declared is false and Open World Assumption (OWA) for those considering that there may be facts that cannot be deduced from what is specified and that can be true. As [1], we consider that accurate modeling of the knowledge of engineering domains, to which we are interested, must be done under the CWA assumption. Indeed, this assumption improves the formal validation of the consistency of system’s specifications with respect to domain properties. Moreover, systems of interest to us are so critical that no assertion should be assumed to be true until consensus is reached on its veracity. Similarly, we also advocate strong typing [1], because our domain models are made in order to complete Event-B models for the system specifications to be formally validated.

Several ontology formalisms exist. The main ones are OWL (Ontology Web Language) [17], PLIB (Part LIBrary) [15] and F-Logic (Frame Logic) [8]. A summary
of the similarities and differences between these ontology modeling formalisms is presented through Table 1.

| Characteristics                     | OWL | PLIB | F-Logic |
|-------------------------------------|-----|------|---------|
| Modularity                          | total | partial | total   |
| CWA vs OWA                          | OWA | CWA | CWA      |
| Inheritance                         | multiple | simple | multiple |
| Typing                              | weak | strong (any element belongs to one and only one type) | weak |
| Contextualization of a property (parameterized attributes) | - | + | + |
| Different views for an element      | - | + | - |
| Graphic representation              | + | - | - |
| Domain Knowledge (static vs dynamic) | static | static | static |

- **PLIB, OWL and F-Logic** implement referencing mechanisms between ontologies. **PLIB** supports partial import: a class of an ontology $A$ can extend a class of an ontology $B$ and explicitly specify the properties it wishes to inherit. Moreover, if nothing is specified, no property will be imported. On the other hand, **OWL** and **F-Logic** use the total import: when an ontology $A$ refers to an ontology $B$, all the elements of $B$ are accessible within $A$.

- **PLIB and F-Logic** use the CWA assumption for constraint verification, **OWL** uses the OWA assumption.

- **OWL and F-Logic** implement multiple inheritance and instantiation. **PLIB** implements simple inheritance and instantiation. On the other hand, with the is case of relation, a **PLIB** class can be a case of several other classes, each class bringing some specific properties.

- **PLIB** is strongly typed (any element belongs to one and only one type), which is not the case for **OWL** and **F-Logic**.

- **PLIB and F-Logic** allow the definition of parameterized attributes using context parameters, which is not possible with **OWL**.

- **PLIB** allows the association of several representations or view points with a concept, which is not possible with neither **OWL** nor **F-Logic**.

- **OWL, PLIB and F-Logic** are focused only on modeling of static domain knowledge. It is for example impossible to specify that the localization of a vehicle can change dynamically while its brand can’t.

We can observe, as stated in [20], that “unfortunately, all the studied formalisms emphasize more on modeling static domain knowledge”. None of these formalisms allows to specify that a knowledge described must remain unchanged or that it is likely to be updated. Moreover, the construction of an **OWL** ontology is done under the OWA assumption and **PLIB** does not allow the specification of rules allowing to deduce new facts from existing ones. Finally, **F-Logic** as **OWL** are weakly typed.
4 Our Approach for Domain Modeling

We have chosen to represent domain knowledge using ontologies since they are semantically richer and therefore allow a more explicit representation of domain characteristics. Thus, in this part, we propose a metamodel, based on that of OWL and PLIB and conforming to the CWA assumption for the representation of the domain of a system whose requirements are captured using the SysML/KAOS method. Our formalism make the Unique Name Assumption (UNA) [1] : the name of an element is sufficient to uniquely identify it among all the others within a domain model. Furthermore, our metamodel is designed to allow the specification of knowledge that is likely to evolve over time. We have identified two graphical syntaxes for the representation of ontologies : the syntax proposed by OntoGraph [6] and the syntax proposed by OWLGred [19]. The OntoGraph syntax is the one used in [11]. Unfortunately, it does not allow the representation of some domain model elements such as attributes or cardinalities. For our case study, we have thus decided to use the OWLGred syntax.

We present through Figures 2, 3, 5 and 6 the main part of the metamodel associated with our domain modeling approach, knowing that yellow elements are those having an equivalent in OWL metamodel and that red ones are those that we have either inserted or customized. Furthermore, some constraints and associations, such as the parentConcept association, have been extracted from the PLIB metamodel. Due to space consideration, we will not highlight all the elements and constraints of the metamodel. Figures 4, 8 and 10 represent respectively the domain model associated to the root level of the SysML/KAOS goal diagram illustrated through Figure 1, that associated with the second level of refinement and that associated with the first one. The domain model associated to the goal diagram root level is named "untitled-ontology-52", the one associated to the first refinement level is named "untitled-ontology-53" and the one associated to the second refinement level is named "untitled-ontology-54".

4.1 Concepts and Individuals, Data Sets and Data Values

The central notion is the notion of Concept which represents a group of individuals sharing common characteristics (Fig. 2). A concept can be declared variable (isVariable=true) when the set of its individuals is likely to be updated through addition or deletion of individuals. Otherwise, it is considered to be constant (isVariable=false). A concept may be associated with another, known as its parent concept, through the parentConcept association, from which it inherits properties. As a result, any individual of the child concept is also an individual of the parent concept.

In untitled-ontology-52 (Fig. 4), a Vehicle is modeled as an instance of Concept named "Vehicle" and its localization is represented through an instance of Concept named "Localization". For readability purposes, we have decided to represent the isVariable attribute only when it is set to true. Since it is possible to dynamically add
or remove localizations, the attribute isVariable of Localization is set to true, which is represented by the stereotype «isVariable». Since the system is designed to control a single vehicle, it is not possible to dynamically add new ones. The involved vehicle is thus modeled as an instance of Individual named "v1" having Vehicle as type.

An instance of DataSet is used to group instances of DataValue having the same type (Fig. 3). Default DataSets are INTEGER, NATURAL for positive integers, FLOAT, STRING or BOOL for booleans. The most basic way to build an instance of DataSet is by listing its elements. This can be done through the DataSet specialization called EnumeratedDataSet.

### 4.2 Relations and Attributes

The notion of Relation is used to capture links between concepts (Fig. 5) and the notion of Attribute links between concepts and data sets (Fig. 6). A relation (Fig. 5) or an attribute (Fig. 6) can be declared variable if the list of maplets related to it is likely to change over time. Otherwise, it is considered to be constant. The association between a relation and a concept is characterized by the cardinality:
DomainCardinality and RangeCardinality (Fig. 5). Each instance of DomainCardinality (respectively RangeCardinality) makes it possible to define, for an instance of Relation re, the minimum and maximum limits of the number of instances of Individual, having the domain (respectively range) of re as type, that can be put in relation with one instance of Individual, having the range (respectively domain) of re as type. The following constraint is associated with these limits: \( (\text{minCardinality} \geq 0) \land (\text{maxCardinality} = \text{null} \lor \text{maxCardinality} \geq \text{minCardinality}) \), knowing that if
Fig. 5  Second part of the metamodel associated with domain modeling

maxCardinality = null, then the maximum limit is infinity. Instances of RelationMaplet are used to define associations between instances of Individual through instances of Relation. In an identical manner, instances of AttributeMaplet are used to define associations between instances of Individual and instances of DataValue through instances of Attribute.

Optional characteristics can be specified for a relation (Fig. 5): transitive (isTransitive, default false), symmetrical (isSymmetric, default false), asymmetrical (isASymmetric, default false), reflexive (isReflexive, default false) or irreflexive (isIrreflexive, default false). It is said to be transitive (isTransitive=true) when the relation of an individual x with an individual y which is in turn in relation to z results in the
Fig. 6 Third part of the metamodel associated with domain modeling

relation of $x$ and $z$. It is said to be symmetric when the relation between an individual $x$ and an individual $y$ results in the relation of $y$ to $x$. It is said to be asymmetric when the relation of an individual $x$ with an individual $y$ has the consequence of preventing a possible relation between $y$ and $x$, with the assumption that $x \neq y$. It is said to be reflexive when every individual of the domain is in relation with itself. It is finally said to be irreflexive when it does not authorize any connection of an individual of the domain with itself. Moreover, an attribute can be functional (isFunctional, default true) if it associates to each individual of the domain one and only one data value of the range.

For readability purposes, we have decided to remove optional characteristics representation and to represent the isVariable attribute only when it is set to true. In untitled-ontology-52 (Fig. 4), Localization is the domain of two attributes: the latitude modeled as an instance of Attribute named "loc_latitude" and the longitude modeled as an instance of Attribute named "loc_longitude". loc_latitude has, as range, an instance of CustomDataSet named "Latitude" and loc_longitude an instance of CustomDataSet named "Longitude". Since it is possible to dynamically change the localization of a vehicle, the attribute isVariable of loc_latitude and that of loc_longitude are set to true, which is represented by the stereotype «isVariable». The association between an instance of Vehicle and an instance of Localization is represented through an instance of Relation named "estimated_location". Its associated instance of DomainCardinality has 1 as minCardinality and maxCardinality, and its associated instance of RangeCardinality has 0 as minCardinality and 1 as maxCardinality.
Fig. 7 Fifth part of the metamodel associated with domain modeling
4.3 Functions and Predicates

The notion of `DataFunction` (Fig. 3) makes it possible to define operations which allow to determine data values at the output of a set of processes on some input data values. At each tuple of data values of the domain, the `data function` assigns a tuple of data values of the range, and this assignment cannot be changed dynamically.

**Example:** We can define an instance of `DataFunction` named "multiply" to produce, given two instances of `INTEGER x` and `y`, the instance of `INTEGER` representing `x * y`.

On the other side, the notion of `Predicate` (Fig. 2) is used to represent constraints between different elements of the domain model in the form of Horn clauses: each predicate has a body which represents its antecedent and a head which represents its consequent, body and head designating conjunctions of atoms. A typing atom is used to define the type of a term: `ConceptAtom` for individuals and `DataSetAtom` for data values (Fig. 7). An association atom is used to define associations between terms: `RelationshipAtom` for the connection of two terms through a relation, `AttributeAtom` for the connection of two terms through an attribute and `DataFunctionAtom` for the connection of terms through a data function (Fig. 7). For each case, the types of terms must correspond to the domains/ranges of the considered link. A comparison atom is used to define comparison relationships between terms: `EqualityAtom` for equality and `InequalityAtom` for difference (Fig. 7). Built in atoms are some specialized atoms,
characterized by identifiers captured through the `AtomType` enumeration, and used for the representation of particular constraints between several terms (Fig. 7). For example, an arithmetic constraint between several integer data values.

In `untitled-ontology-54` (Fig. 8), the constraint "a GPS is more precise than a Wi-Fi" is translated into an instance of `Predicate` represented through formula 1. If an instance of `Term`, named "x", has `Wifi` as its type, has `px` as its `precision` and an instance of `Term`, named "y", has `Gps` as its type, has `py` as its `precision`, then `py > px`.

\[
greaterThan(?py, ?px) \leftarrow \text{Wifi}(?x) \land \text{precision}(?x, ?px) \land \text{Gps}(?y) \land \text{precision}(?y, ?py)
\] (1)

`Predicates` can be used to parameterize relations or attributes in order to define dependent associations. For example, knowing that the resistance of a material depends on the temperature of the medium, resistance and temperature attributes are dependent. `GluingInvariant` (Fig. 2), specialization of `Predicate`, is used to represent links between variables elements defined within a domain model and those appearing in more abstract domain models, transitively linked to it through the `parent` association. Gluing invariants are extremely important because they capture relationships between abstract and concrete variables during refinement that are used to demonstrate proof obligations.

### 4.4 Domain Model and Goal Model

Each domain model is associated with a level of refinement of the `SysML/KAOS` goal diagram and is likely to have as its parent, through the `parent` association, another domain model (Fig. 2). This allows the child domain model to access and extend some elements defined in the parent domain model. It should be noted that the parent domain model must be associated with the refinement level of the ` SysML/KAOS` goal diagram directly above the refinement level to which the child domain model is associated.

`untitled-ontology-53` (Fig. 10) has `untitled-ontology-52` (Fig. 4) as parent and defines new concepts and relationships. Each reused element is annotated with the name of its domain model. This third abstraction level represents child concepts of `SubComponent` and `Sensor`. A
subcomponent is either a GPS, represented through an instance of Concept named "Gps", or a Wi-Fi, represented through an instance of Concept named "Wifi". A sensor is either an accelerometer, represented through an instance of Concept named "Accelerometer", or a speed sensor, represented through an instance of Concept named "SpeedSensor". Finally, \( v1 \) is associated to an instance of Individual of type Gps named "g1" and to an instance of Individual of type Wifi named "w1" through vehicle_subcomponents, an instance of Relation introduced in untitled-ontology-53. It is also associated to a speed sensor called \( s1 \) and to an accelerometer called \( a1 \).

In order to be able to be used in the setting up of large complex systems, SysML/KAOS allows the refinement of a leaf of a goal diagram in another diagram having this goal as root. For example, in Figure 9, the goal \( G3 \), which is a leaf of the first goal diagram, is the root of the second one. When this happens, we associate to the most abstract level of the new goal diagram the domain model associated with the most concrete level of the previous goal diagram as represented in Figure 9: Domain Model 2, which is the domain model associated to the most concrete level of the first diagram, is also the domain model associated to the root of the second one.

![Fig. 9 Management of the partitioning of a SysML/KAOS goal model](image)

5 Conclusion

In this paper, we have drawn up the state of the art related to domain modeling in requirements engineering. After positioning ourselves as to the existing, we have presented our domain modeling method consisting in representing domain knowledge using an ontology modeling formalism for which a metamodel has been defined. Our approach has been illustrated through a case study dealing with a Cycab localization component specification.

Work in progress is aimed at developing mechanisms for the explicitness of SysML/KAOS domain models semantics in Event-B and at integrating our approach within the open-source platform Openflexo [14].

Acknowledgements This work is carried out within the framework of the FORMOSE project [3] funded by the French National Research Agency (ANR).
Fig. 10  *untitled-ontology-53*: ontology associated to the first level of refinement

References

1. Ameur, Y.A., Baron, M., Bellatreche, L., Jean, S., Sardet, E.: Ontologies in engineering: the OntoDB/OntoQL platform. Soft Comput. 21(2), 369–389 (2017)
2. ANR-06-SETIN-017: TACOS ANR project (2017). URL [http://tacos.loria.fr/](http://tacos.loria.fr/)
3. ANR-14-CE28-0009: Formose ANR project (2017). URL [http://formose.lacl.fr/](http://formose.lacl.fr/)
4. Dermeval, D., Vilela, J., Bittencourt, I.L., Castro, J., Isotani, S., da S. Brito, P.H., Silva, A.: Applications of ontologies in requirements engineering: a systematic review of the literature. Requir. Eng. 21(4), 405–437 (2016)
5. Devedzic, V.: Knowledge modeling - state of the art. Integrated Computer-Aided Engineering 8(3), 257–281 (2001)
6. Falconer, S.: Protégé - ontograph (2010). URL [http://protegewiki.stanford.edu/wiki/OntoGraf](http://protegewiki.stanford.edu/wiki/OntoGraf)
7. Gnaho, C., Semmak, F.: Une extension SysML pour l’ingénierie des exigences dirigée par les buts. In: 28e Congrès INFORSID, France, pp. 277–292 (2010)
8. Kifer, M., Lausen, G.: F-logic: A higher-order language for reasoning about objects, inheritance, and scheme. In: Proceedings of the 1989 ACM SIGMOD, pp. 134–146. ACM Press (1989)
9. Kitamura, M., Hasegawa, R., Kaiya, H., Saeki, M.: An integrated tool for supporting ontology driven requirements elicitation. In: J. Filipe, B. Shishkov, M. Helfert (eds.) ICSoft 2007, Volume SE, pp. 73–80. INSTICC Press (2007)
10. van Lamsweerde, A.: Requirements Engineering - From System Goals to UML Models to Software Specifications, Wiley (2009)
11. Mammar, A., Lalau, R.: On the use of domain and system knowledge modeling in goal-based Event-B specifications. In: ISoLA 2016, Lecture Notes in Computer Science, vol. 9952, pp. 325–339 (2016)
12. McUmber, W.E., Cheng, B.H.C.: A general framework for formalizing UML with formal languages. In: ICSE 2001, pp. 433–442. IEEE Computer Society (2001)
13. Nguyen, T.H., Vo, B.Q., Lumpe, M., Grundy, J.: KBRE: a framework for knowledge-based requirements engineering. Software Quality Journal 22(1), 87–119 (2014)
14. Openflexo: Openflexo project (2015). URL [http://www.openflexo.org](http://www.openflexo.org)
15. Pierra, G.: The PLIB ontology-based approach to data integration. In: IFIP 18th World Computer Congress, *IFIP*, vol. 156, pp. 13–18. Kluwer/Springer (2004)
16. Sekhavat, S., Valadez, J.H.: The Cycab robot: a differentially flat system. In: IROS 2000, pp. 312–317. IEEE (2000)
17. Sengupta, K., Hitzler, P.: Web ontology language (OWL). In: Encyclopedia of Social Network Analysis and Mining, pp. 2374–2378 (2014)
18. Shibaoka, M., Kaiya, H., Saeki, M.: GOORE : Goal-oriented and ontology driven requirements elicitation method. In: ER 2007 Workshops, *Lecture Notes in Computer Science*, vol. 4802, pp. 225–234. Springer (2007)
19. UL, I.: Owlged home (2017). URL http://owlged.lumii.lv/
20. Zong-yong, L., Zhi-xu, W., Ai-hui, Z., Yong, X.: The domain ontology and domain rules based requirements model checking. *International Journal of Software Engineering and Its Applications* 1(1), 89–100 (2007)