Using seal trajectories in biological early warning system for real-time zone tracking

Rouaa Wannous[1], **Jamal Malki**[1], **Alain Bouju**[1], **Cécile Vincent**[2]

1. *Laboratoire L3i, Pôle Sciences & Technologie, Avenue Michel Crépeau, 17042 La Rochelle, France*
   rwannous,jmalki,abouju@univ-lr.fr
2. *Centre d'Etudes Biologiques de Chizé, CNRS UMR 7372, 79360 Villiers-en-Bois, France*
   cvincent@univ-lr.fr

ABSTRACT. *Early warning systems were interested in captured data of mobile objects. From the 2000s, a new generation of data capture equipment arrives. These capture devices rise large scale trajectory data. How early warning systems can integrate these masses of data? How they can give real-time answers to users queries? In this paper, we present an ontological approach to model the trajectory. The trajectory's domain knowledge are expressed as rules used by the ontological inference mechanism. We show the important complexity of the inference and we propose optimizations. We evaluate our contributions over real data.*

RÉSUMÉ. *Les systèmes d'alertes rapides se sont intéressé aux données capturées en particulier celles des objets mobiles. Depuis le début des années 2000, de nouveaux dispositifs de capture sont conçus qui sont capables de restituer de grands volumes de données, appelés trajectoires. Comment les systèmes d'alertes rapides peuvent-ils intégrer ces masses de données ? Comment répondre en temps réel aux requêtes des utilisateurs ? Dans ce travail, nous présentons une approche ontologique pour modéliser la trajectoire. Les connaissances du domaine de la trajectoire sont exprimées sous forme de règles qui alimentent le mécanisme d'inférence ontologique. Nous montrons l'importante complexité de l'inférence et nous proposons des optimisations. Nous évaluons les contributions sur des données réelles.*

KEYWORDS: *early warning system, trajectory ontology modeling, ontology inference, domain rules, temporal rules, data filter algorithm.*

MOTS-CLÉS : *système d'alerte raide, modélisation ontologique des trajectoires, ontologie et inférence, règles de domaine et temporelles, algorithme de filtrage de données.*

DOI:10.3166/ISI.21.4.83-104

1. Introduction

Recent advances in communication and sensor technology have catalyzed progress in early warning system for zone tracking. A global early warning system is needed to inform us of pending threats. The basic idea behind early warning is that the earlier and more accurately we are able to predict short and long term potential risks associated with natural and human induced hazards, the more likely we will be able to manage and mitigate a disaster's impact on society, economies, and environment. Early warning (UNEP, 2012) is "the provision of timely and effective information, through identified institutions, that allows individuals exposed to hazard to take action to avoid or reduce their risk and prepare for effective response".

Research has focused on a biological early warning system over the last few decades (Butterworth, Gonsebatt, 2001; Green *et al.*, 2003). In order to apply the system to regional environmental conditions and achieve different monitoring aims, biological early warning system has been diversified using various test organisms such as water fleas, mussels, algae, and fish (Baldwin *et al.*, 1994; Benecke, Schmidt, 1982; Borcherding, Jantz, 1997; Hendriks, Stouten, 1993). Within the framework of defining and building a warning biological information system, several animals are used: mammal, bird, herp, invert, etc. Trajectories of these species are captured and analyzed over long periods. Thanks to the collected data and using other external data sources, we are now able to identify precise states of these animals in their natural environment. Different types of behavioral parameters have been developed for various kinds of monitoring systems related to their state.

In this paper, we present our approach for integrating trajectories of marine mammal, namely seals, in an early warning tracking system. The raw data captured, commonly called trajectories, traces animals from a departure point to a destination point as data sequences (sample points captured, time of the capture). Trajectory data are captured by sensors included in a tag glued to the fur of the animal behind the head. The captured trajectories consist of spatial, temporal and spatio-temporal data. Trajectories data can also contain some extra-data. These datasets are organized into sequences. Every sequence, mapped to a temporal interval, characterizes a defined state of the animal. In our application, we consider three main states of a seal: *hauling out, diving* and *cruising*. Every state is related to a seal's activity. For example, a foraging activity occurs during the state diving.

To detect the appropriate or protected zone used by a seal, this study develops an early warning system which integrates an alarm rule with a seal trajectory model. For that, we detect the seals' foraging areas in order to assess the interactions with the human fisheries activities. Our aim is to quickly and automatically identify those important (foraging) areas from the seals trajectories. First of all, we need a trajectory model. In our previous work (Wannous *et al.*, 2013b), we define a trajectory ontology model taking into account domain, spatial and temporal data. Using the ontological rules associated with this model, we compute the inference over these data. The experimental results addressed time computation and space storage problems of the ontology inference. Then, we proposed some solutions to reduce the inference complexity by defining time restrictions in the inference passes refinements in (Wannous *et al.*, 2013a). These later studies focus mainly on the term of time computation.

In the present paper, we continue studying the ontological inference complexity, specially in terms of inference space storage complexity. We propose two-tier inference filters on trajectory data. In other words, two distinct operations are performed to enhance the inference: primary and secondary filter operations. The primary filter is applied to the captured data with the consideration of domain constraints. The primary filter allows fast selection of the analyzed data to pass along to the secondary filter. The latter computes the inference over the data output of the primary filter.

This paper is organized as follows. Section 2 summarizes recent work related to early warning and monitoring systems with a focus on those based on trajectory data. In this context, we focus on approaches that define data models taking into account low level and semantic aspects. Section 3 introduces our approach and illustrates an overview of our domain data model called "trajectory ontology". This trajectory ontology defines temporal concepts mapped to W3C OWL-Time ontology (Jerry, Feng, 2004) in Section 4. Sections 5 and 6 detail the trajectory ontology inference and the integrated knowledge. In section 7, we implement the trajectory ontology, the domain ontology rules and the temporal rules. Section 8 addresses the complexity of the ontology inference over the domain and temporal rules. Section 9 introduces an application domain inference refinement. Section 10 evaluates the ontology inference over the proposed refinement and present some result about seal zone tracking for the early warning system. Finally, Section 11 concludes this paper and presents some prospects.

2. Related work

A state of the art analysis and future directions (UNEP, 2012) is given to present the environmental threats, including air quality, impacts of climate variability, severe weather, storms. This report identifies current gaps and needs with the goal of laying out guidelines for developing an early warning system. The aim of this report is to identify current gaps and future needs of early warning systems through the analysis of the state of the art on existing early warning and monitoring systems for environmental hazards.

In (Doong *et al.*, 2012), authors present a study for developing a coastal flooding early warning system (CoFEWs) by integrating existing sea-state monitoring technology, numerical ocean forecasting models, historical database and experiences, as well as computer science. A warning signal is presented when the storm water level that accumulated from astronomical tide, storm surge, and wave-induced run-up exceeds the alarm sea level.

Biological early warning and emergency management support system (Yuan *et al.*, 2009) is given for water pollution accident. This research presents a system which integrates an online water quality monitoring device with a water quality model. The system has been instantiated in Douhe Reservoir. The monitoring device is based on water quality probes and biological sensors which use fish motion as indicator. Another biological early warning system is developed by (Jeong *et al.*, 2014) from swimming behavior of Daphnia magna.

Based on captured data, an early warning system needs a modeling approach to understand and to analyze these data, that we also call trajectories. Recently, several approaches were developed because the access to the captured data became real and easy especially with the advent of open data sharing platforms, like Movebank Data Repository (Crofoot *et al.*, 2015).

Based on birds migration captured data, an approach called a conceptual view on trajectories is introduced by (Spaccapietra *et al.*, 2008). In this approach, trajectories are considered as a set of stops and moves. Each part contains a set of semantic data. Based on this conceptual model, several studies have been proposed. In (Alvares *et al.*, 2007), authors proposed a trajectory data preprocessing method to integrate trajectories with spatial data. Their application concerned daily trips of employees from home to work and back. However, the scope of their paper is limited to the formal definition of semantic trajectories with the space and time without any implementation and evaluation. Researchers in (Yan *et al.*, 2010) proposed a trajectory computing platform which exploits a spatio-semantic trajectory model. A difference is made between the semantic and spatial dimensions in order to provide a data model representation that supports different abstraction levels. Authors present a solution for extracting semantic trajectories from raw data. One of the layers of the presented platform is a data preprocessing layer which cleanses the raw GPS feed, in terms of preliminary tasks such as outliers removal and regression-based smoothing. However, this work did not discuss the computation complexity of the platform going from the raw data step processing to knowledge extraction and finally decision-making.

Based on a space-time ontology and events approach, Boulmakoul et al. (Boulmakoul *et al.*, 2012) proposed a trajectory patterns of moving objects. Important packages of the trajectory patterns are “Space Time Path Domain”, “Activity Domain”, “Observation and Measure Domains” and “Region Of Interest” packages. These packages are then transformed onto a unified moving object trajectory queries expressed in SQL-like relational database language. Queries operations on space and time are performed using simple relational entities and functions. So they seem to rely on a pure SQL-based approach not on semantic queries. This work also did not discuss the evaluation of the proposed approach on real data sets.

In (Boulmakoul *et al.*, 2013), authors gave a brief outline of a scalable data collection framework for the unified moving object trajectories meta model. They gathered different kinds of geographical data based on the unified moving object trajectories' meta-model. The collection framework offers components to collect spatio-temporal data. They test the scalability of the proposed system by a vehicle tracking simulator which generates and simulates spatio-temporal data of different moving objects. Recently, Boulmakoul et al. (Boulmakoul *et al.*, 2015) proposed a trajectory's data model which has advantages of both conceptual and ontological space-time. So they extend the model with new patterns as the space-time path to describe activities of the moving object and the composite region of interest. The case study is presented for tracking travelers at the airport.

3. Modeling approach

3.1. Design and methodology

Our work is based on moving objects trajectories. This requires a trajectory data model and a moving object model. Moreover, to enrich data with knowledge, a semantic model should be taken into consideration. Therefore, we need a generic model to consider the trajectory, moving object and semantic models simultaneously as shown in Figure 1. The semantic trajectory model can consume captured data of trajectories and other external data as shown in Figure 1 link (1). These data are related to an application domain. This requires an application domain trajectory model which consists of domain model, as shown in Figure 1 link (2). The latter will support semantics related to users' needs. In the domain model, we also find the necessary semantics related to the real moving object, its trajectories, its activities and others. This semantics is often designed by a domain expert. In general, considering various facets of data involves that the semantic trajectory model must be extended by other models: application domain, temporal and spatial models. Then, the main issue is to build and design the semantic trajectory model with its required components.

The semantic trajectory modeling approach is tightly related to the problem of a semantic gap between this model and raw data. Link (1) in Figure 1 presents this gap. Moreover, our approach involves multiple models and then must establish semantic mappings among them, to ensure interoperability. In Figure 1, links (2) and (3) match the domain, temporal and spatial models with the semantic trajectory model. This matching extends the capabilities of our approach. For more efficient semantic capabilities, we want to annotate the data with domain, temporal and spatial knowledge. These knowledge are defined by experts representing users' needs. Annotating data with these knowledge could be done automatically or manually. We cannot use a man-

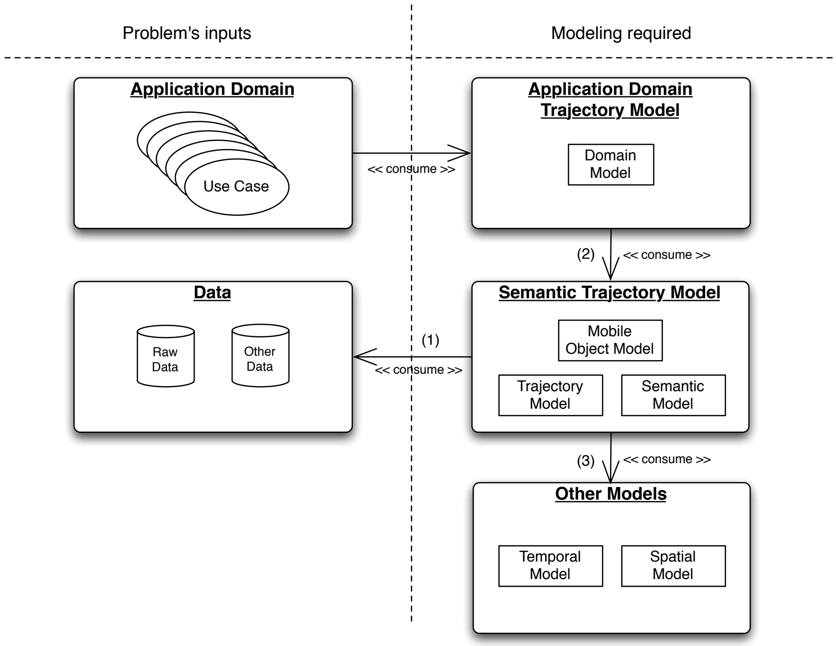


Figure 1. Problem and its modeling required

ual annotation over huge data. Therefore, we choose an automatic annotation which can be accomplished by an ontology inference mechanism. This inference mechanism derives new semantics from existing information using additional knowledge. Later in this paper, we will present this inference mechanism as sets of rules.

3.2. Semantic trajectory ontology

In (Malki *et al.*, 2012; Wannous, 2014), we proposed a methodology for modeling trajectory data. This methodology focused on several real cases. For each case, we define a context, data capture, an analysis process of these data, and a domain model. From these models, we define a trajectory pattern also called generic trajectory model, Figure 2. A trajectory is a set of sequences of spatio-temporal path covered by a moving object and has an activity.

To build the trajectory ontology, we use model transformation techniques introduced by the Model Driven Engineering (MDE) community. For this, we choose an automatic transformation from UML model into a formal OWL ontology. We use transformer tool called uml2owl [1] (Hillairet, 2007). This

1. http://perso.univ-lr.fr/ghillair/projects.html

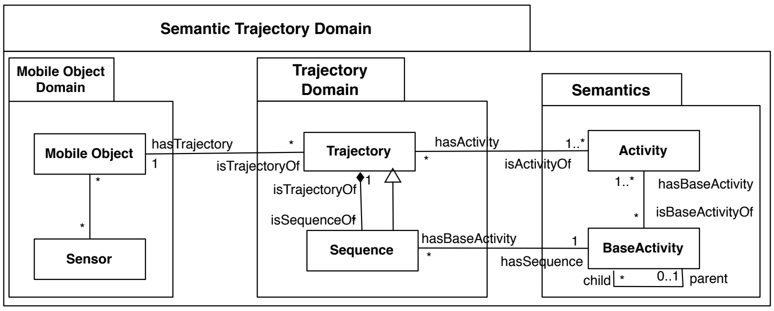


Figure 2. Semantic trajectory modeling approach

transformer, based on the meta-model eCore Eclipse, takes as input a UML model and turns it into OWL-DL ontology. So, we transform the trajectory data model (Figure 2) to an OWL ontology, named `owlSemanticTrajectory`. Figure 3 presents the declarative part of this ontology. It contains three parts: mobile object, trajectory and semantic ontologies. By definition, a trajectory is a set of spatio-temporal concepts. Spatial and temporal models can be reused to enrich description of the concepts in the trajectory ontology to represent their spatial and temporal properties. Table 1 gives a dictionary of the main concepts of the sematic trajectory ontology. Table 2 explains the relationships between concepts in the semantic trajectory ontology.

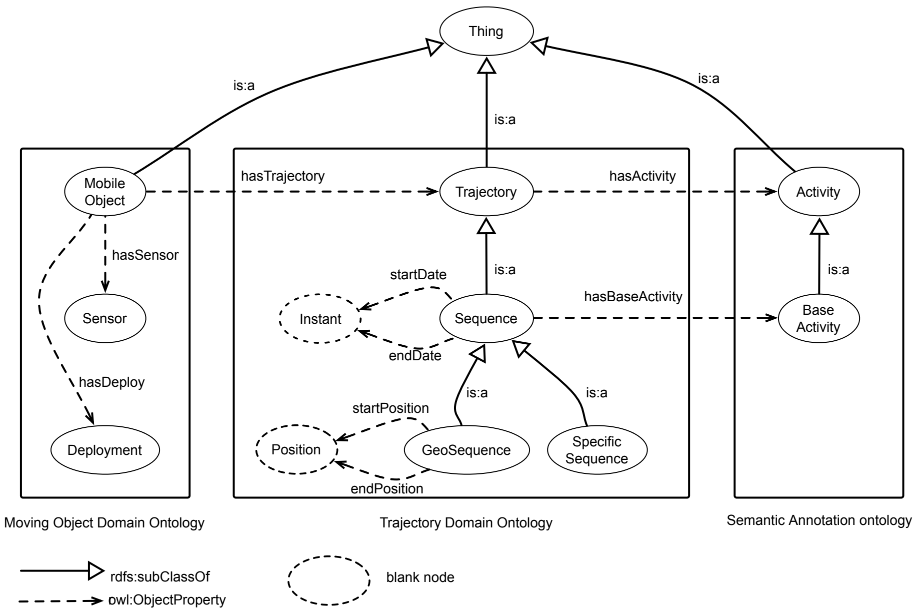


Figure 3. A view of the semantic trajectory ontology `owlSemanticTrajectory`

Table 1. Dictionary classes of the semantic trajectory domain

Classe	**Description**
Trajectory	logical form to represent sets of sequences
Sequence	spatio-temporal interval representing a capture
GeoSequence	spatial part of a sequence
Specific Sequence	metadata part of a sequence
Activity	mobile object's activity in a sequence
Mobile object	the moving object equipped with a sensor

Table 2. Relationships between concepts in the semantic trajectory domain

Classe	**Description**
hasActivity/hasBaseActivity	an object property to the activity of a trajectory/sequence
startPosition, endPosition	the capture position of a geosequence
startDate, endDatee	the capture time of a sequence
hasTrajectory	the trajectory of a mobile object

4. Time ontology

The seal trajectory ontology includes concepts that can be considered as temporal. The concept `Sequence` is a temporal interval. To integrate temporal concepts and relationships in the seal trajectory ontology, we choose a mapping approach between our ontology and the OWL-Time [2] ontology (Jerry, Feng, 2004) developed by the World Wide Web Consortium (W3C). This mapping is detailed in our previous work (Wannous *et al.*, 2013b). An extract of the declarative part of this ontology is shown in Figure 4 described in detail in (Jerry, Feng, 2004). We are mainly interested in the `ProperInterval` concept and its two properties `hasBeginning` and `hasEnd`.

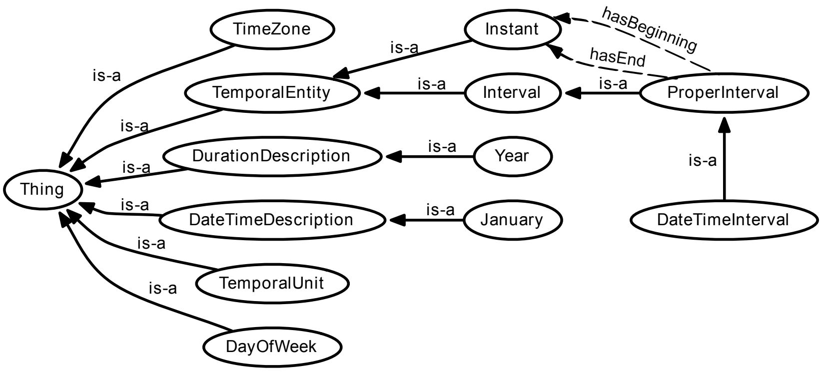


Figure 4. A view of the OWL-Time ontology

2. http://www.w3.org/2006/time

5. Trajectory ontology inference

Inference is the ability to make logical deductions based on ontology model, rules and optionally individuals. It derives new knowledge based on rules. A rule's definition, Figure 5, has an antecedent, filters and a consequent. If knowledge is represented using RDF triples, then the antecedent is a set of triples, filters apply restrictions, and finally consequent is a new derived triple.

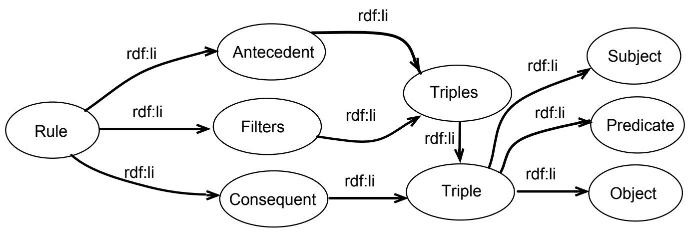


Figure 5. Rule's definition

In the present work, we consider two kinds of inference:

1. Inference using standard rules: Our semantic trajectory ontology is based on RDF, RDFS and OWL constructs. Inference mechanism associates with each construct a rule. The results sets are called standard rules. An example of standard rules is OWLPrime in Oracle RDF triple store (Oracle, 2009).

2. Inference using temporal rules: Our semantic trajectory ontology uses temporal relationships as defined by Allen's algebra (Allen, 1983). Each relationship is defined as a rule such as: `intervalAfter`, `intervalBefore`, `intervalDuring`, etc.

6. Trajectory ontology inference using domain rules

Our application domain is seals' trajectories, where a seal is considered as a mobile object. The captured data comes from the LIENSs laboratory[3] in collaboration with SMRU[4]. We consider three main states of a seal : `Dive`, `Haulout` and `Cruise`. Every state is related to a seal's activity, like `Resting`, `Traveling` and `Foraging`.

The captured data can also contain some meta-data called `CTD` (Conductivity-Temperature-Depth) about the marine environment such as water conductivity, temperature and pressure. Starting from our semantic trajectory ontology `owlSemanticTrajectory` we define the seal trajectory ontology, named `owlSealTrajectory`, Figure 6. Formally, each activity is declared in the ontology and associated to a domain rule.

3. Lab. CNRS/University of La Rochelle - http://lienss.univ-larochelle.fr
4. SMRU: Sea Mammal Research Unit - http://www.smru.st-and.ac.uk

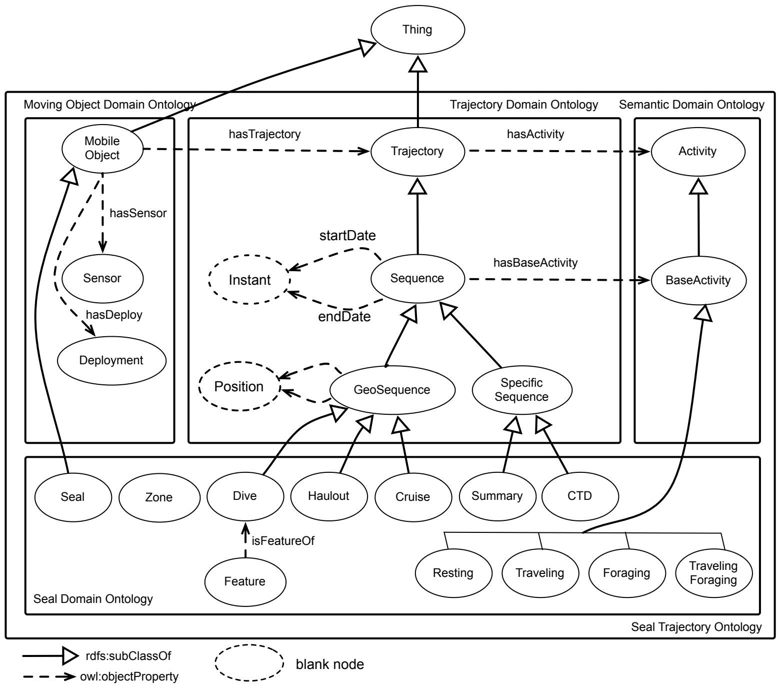


Figure 6. Overview of the seal trajectory ontology with their activities

7. Implementation

Our implementation framework uses Oracle RDF triple store (Oracle, 2009). Based on a graph data model, RDF triples are persisted, indexed and queried, like other object-relational data. In this framework, we create the following models and rulebases (a set of rules):

– `owlTrajectory`, `owlTime` and `owlSealTrajectory`: declarative part of the trajectory, time and seal ontologies;

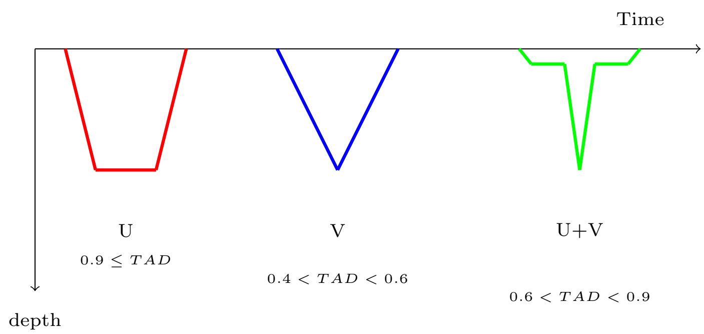
– `OWLPrime`: rulebase of the standard rules;

– `Time_Rules`: a rulebase of the temporal holding the interval temporal relationships. The declarative part of the `intervalAfter_rule` is presented in Figure 7 based on operations defined in the `TM_RelativePosition` table of the ISO/TC 211 specification about the temporal schema (ISO/TC_211, 2002).

– `Seal_Rules`: a rulebase of the seal rules. According to the domain expert, there is a correlation between the geometrical shape of dives and activities. To classify geometric shapes of dives, the `TAD` (Fedak *et al.*, 2001) index is computed over a set of data. For this classification, we can distinguish three patterns:

- dive shaped V: if $0 <= TAD < 0.7$

- dive shaped U+V: if $0.7 <= TAD < 0.9$
- dive shaped U: if $0.9 <= TAD < 1$



According to the domain expert, we take into consideration different parameters to define the seal activities. The parameters are the geometrical shape of dives (`TAD`), the maximum dive depth and surface ratio which is the ratio between surface duration and dive duration. The decision Table 3 summarizes conditions of the IF parts of rules associated with activities. Based on this table, Figure 8 gives an example of rule definition, `foraging_rule`, in the system. Detail of the activities:

– `Resting` is when a seal is sleeping at the sea bottom with the TAD higher than 0.9. The surface duration after the dive state should be quite high so that seals have enough time to breath before another sleep under water;

– `Traveling` could be in any dive depth deeper than 3 meters, but the TAD should be lower than 0.7 because the seal does not need to spend a lot of time at the maximum depth. The surface duration does not make any difference in this case;

– `Foraging` is when the dive depth is deeper than 3 meters. The TAD however should be high (>0.9) because the grey seal is a benthic forager, which means it is feeding on fish located on or close to the sea bottom (i.e at the maximum depth available). Also the surface duration is short because the seal wants to go back quickly to look for more fish;

– `TravelingForaging` is when the dive depth is deeper than 3 meters. The TAD however should be higher than (>0.7) and smaller than (>0.9). Also the surface duration is short because the seal wants to go back quickly to look for more fish.

In our framework, inference mechanism creates a rule index, Figure 9. A rule index (entailment) is an object containing pre-computed triples from applying a specified set of rulebases to a specified set of models. If a graph query refers to any rulebases, a rule index must exist for each rulebase-model combination in the query. The `USER_RULES=T` option is required while applying user-defined