



**HAL**  
open science

# Semantic Spatial Representation: a unique representation of an environment based on an ontology for robotic applications

Guillaume Sarthou, Rachid Alami, Aurélie Clodic

## ► To cite this version:

Guillaume Sarthou, Rachid Alami, Aurélie Clodic. Semantic Spatial Representation: a unique representation of an environment based on an ontology for robotic applications. Combined Workshop on Spatial Language Understanding (SpLU) and Grounded Communication for Robotics (RoboNLP), Jun 2019, Minneapolis, United States. hal-02177140

**HAL Id: hal-02177140**

**<https://hal.science/hal-02177140>**

Submitted on 8 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Semantic Spatial Representation: a unique representation of an environment based on an ontology for robotic applications

**Guillaume Sarthou**  
LAAS-CNRS, Université  
de Toulouse, CNRS,  
Toulouse, France  
gsarthou@laas.fr

**Aurélie Clodic**  
LAAS-CNRS, Université  
de Toulouse, CNRS,  
Toulouse, France  
aclodic@laas.fr

**Rachid Alami**  
LAAS-CNRS, Université  
de Toulouse, CNRS,  
Toulouse, France  
alami@laas.fr

## Abstract

It is important, for human-robot interaction, to endow the robot with the knowledge necessary to understand human needs and to be able to respond to them. We present a formalized and unified representation for indoor environments using an ontology devised for a route description task in which a robot must provide explanations to a person. We show that this representation can be used to choose a route to explain to a human as well as to verbalize it using a route perspective. Based on ontology, this representation has a strong possibility of evolution to adapt to many other applications. With it, we get the semantics of the environment elements while keeping a description of the known connectivity of the environment. This representation and the illustration algorithms, to find and verbalize a route, have been tested in two environments of different scales.

## 1 Introduction

Asking one's way, when one does not know exactly where one's destination is, is something we all did. Just as we have all responded to such a request from a lost person. This is the heart of the road description task. This task, which seems so natural to us in a human-human context, requires in fact a set of knowledge (e.g. on the place, on the possible points of reference in this particular environment) and "good practices" (e.g. to give a path easy to follow if possible) that need to be modeled if we want to implement it on a robot. This paper presents a robotics application of our system but it would be possible to use it in other applications such as virtual agent.

This route description task is an interesting application case through the variety of the needed information (e.g. type of elements, place topology, names of the elements in natural language). It has been well studied in the field of human-robot interaction. Robot guides have already been deployed

in shopping centers (Okuno et al., 2009), museums (Burgard et al., 1998; Clodic et al., 2006; Siegwart et al., 2003) or airports (Triebel et al., 2016). However, using metrical (Thrun, 2008), topological (Morales Saiki et al., 2011), semantic representations, or trying to mix them together (Satake et al., 2015b) (Chrastil and Warren, 2014) (Zender et al., 2008), it is difficult to have a uniform way to represent the environment. In addition, it is difficult to have a representation which allows to calculate a route and at the same time to express it to the human with whom the robot interacts, because this requires data of different types. Our aim is to propose a single and standardized representation of an environment which can be used to choose the appropriate route to be explained to the human and at the same time to verbalize it using a route perspective. The purpose of this paper is not to be applied to a guiding task in which a mobile robot accompanies the human to his final destination but to explain to a human how to reach it. Consequently, we will not talk here about a metrical representation like the one that can be used to navigate in the environment (Thrun, 2008) or to negotiate it use in (Skarzynski et al., 2017).

Route perspective means essentially to navigate mentally in order to verbalize the path to follow but also to facilitate understanding and memorizing instructions. The route perspective opposes the survey perspective which is a top view with landmarks and paths printed on a map. Morales et al. (Morales et al., 2015) indicate that naming parts of a geometric map does not leave the opportunity to compute such perspective. As in (Satake et al., 2015a), we have chosen to develop our representation with an ontology as it allows to reason about the meaning of words and thus improve the understanding of human demands. In addition, we propose a way to merge the topological representation into the semantic representation

(the ontology) to get the meaning of the environment elements while keeping a description of the connectivity of the elements of the environment. We propose to name it semantic spatial representation (SSR). With this, we are able to develop the two features presented in (Mallot and Basten, 2009) for the route description task, which consist of selecting a sequence of places leading to the objective and managing the declarative knowledge to choose the right action to explain at each point of the sequence. Based on the principles of topological description, although represented semantically, we are able to compute multiple routes and new detours for the same objective in contrast with a route knowledge, which maps a predefined route to a given request. Thanks to this capacity and to the semantic knowledge of the environments available in the representation, it is also possible to provide the most relevant route to a user according to his preferences and capabilities. A basic example would be that we will never recommend a path with stairs to a mobility impaired person. More than the extension of the spatial semantic hierarchy (SSH) (Kuipers, 2000) allowing the representation of the environment, we present here an algorithm to choose the routes and another one to generate an explanation sentence. Both algorithms are based solely on the knowledge provided by the SSR.

Regarding the representation of the environment generally used in order to find an itinerary, we have first to analyse GNSS road navigation systems. In (Liu, 1997) or (Cao and Krumm, 2009), we find the same principle of a topological network representing the roads with semantic information attached to each of them. This type of representation seems logical regarding the performance required for such systems operating in very large areas. However, GNSS road navigation systems must respond only to this unique task of finding a path when a robot is expected to be able to answer to various tasks. This is why we have developed and implemented a representation that can be used more widely while still allowing the search for routes.

This paper focuses on the presentation of the SSR and on its usability for the route description task. For now, all the ontologies used to test the SSR have been made by hand. However, many recent research work leads to automatically generate a topological representation of an environment

from geometric measurements (e.g. Region Adjacency Graphs (Kuipers et al., 2004), Cell and Portal Graphs (Lefebvre and Hornus, 2003) or hierarchical models (Lorenz et al., 2006), or from natural language (Hemachandra et al., 2014)). We have not done it yet, but our system could benefit from this work to generate a representation of an environment using SSR, which would solve the complexity of creating such a representation by hand.

In order to present our work, we will follow the three cognitive operations needed to generate a spatial discourse (Denis, 1997): (section 2) an activation of an internal representation of the environment; (sections 3 and 4) the planning of a route in the mental representation made previously; (section 5) the formulation of the procedure that the user must perform to achieve the objective. The SSR and the algorithms demonstrating its usability have been tested in two environments of different scales: an emulated mall in our laboratory and a real mall. Results are presented in section 6 for the two environments.

## 2 Environment representation: SSR

In cognitive psychology, Semantic memory refers to the encyclopedic knowledge of words associated to their meanings. Some authors have proposed a model of this semantic memory as being a semantic network in which each concept is linked to others by properties and have designed a computer-implemented model (Collins and Quillian, 1969). This initial model has since been formalized as an ontology (Berners-Lee et al., 2001) and is already widely used in the semantic web.

This model is already used in robotics to obtain a detailed representation of the environment in which robots operate. For example, (Satake et al., 2015a) and (Beetz et al., 2018) use an ontology to represent knowledge about the types of items such as the types of shops (restaurant, fashion store, for example) or the properties of items such as the stores where they are sold.

(Kuipers, 2000) introduced the 'topological level' with SSH (spatial semantic hierarchy) which defines a place, a path and a region and defined several relationships between them. Ontologies are constructed using triplets where two concepts are linked by a property (e.g. *property(concept1, concept2)*), however the Kuipers SSH does not allow such representation due to the use of some quadruplets (e.g.

$along(view, path, dir)$ ) in addition to triplets. To overcome this limitation, we propose a formalisation, that we call Semantic Spatial Representation (SSR) to represent an environment with ontologies (i.e. using triplets).

In this section we present the minimal ontology that constitutes the SSR but it can be extended to represent the knowledge of the types and the properties of the elements while preserving the first use of this model.

## 2.1 Classes

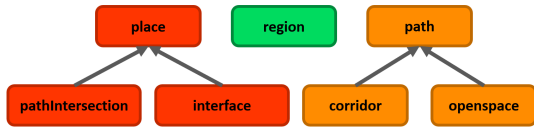


Figure 1: Classes for a representation of the topology of an indoor environment in a semantic description.

*Region*: It represents a two-dimensional area that is a subset of the overall environment. A description of the environment must include at least one region representing the entire environment. Regions are used to reduce the complexity of the routes computation, so we recommend to use several region especially for large scale areas. A basic use of the regions is for multi-storey buildings where each floor should be more naturally considered as a region. Regions can be described as being nested.

*Path*: It is a one dimensional element along which it is possible to move. A path must have a direction.

- *Corridor*: It represents a kind of path with a beginning and an end, for which beginning and end are distinct. The arbitrary direction chosen for a corridor defines the position of its beginning and end. This defines also the right and left of the corridor.
- *Openspace*: It is a kind of path which does not have any begin or end. It can be viewed as a "potato-shaped" describing the outline of an open space. It materializes the possibility of turning the gaze around the room and the fact of not having to go through a defined path to reach one of its points.

*Place*: It represents a point of zero dimension that can represent a physical or symbolic element. It can be extended to represent stores and landmarks in the example of a shopping center.

- *Path intersection*: It represents the connection between only two paths and thus a waypoint to go from one path to another. In the case of a crossing between three paths, three intersections would therefore be described.
- *Interface*: It represents the connection between only two regions and thus a waypoint to move from one region to another. It can be physical, like a door or a staircase, or symbolic like a passage.

The distinction between paths and places is related to the differences between the types of rooms made by (Andresen et al., 2016) where some are used to circulate (corridors) while others have an explicit use to the exclusion of traffic (place).

## 2.2 Properties

Properties are used to express topological relationships such as connections between paths and places or the order of places along a path. All the properties presented here can be extended with their inverse (e.g. *isIn* and *hasIn*) for a more expressive model and thus easier handling.

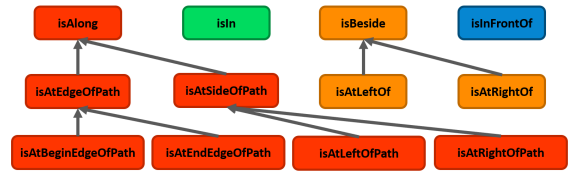


Figure 2: Properties for a representation of the topology of an indoor environment in a semantic description.

$isIn(path/place, region)$ :  $path$  or  $place$  is in  $region$ .

$isAlong(place, path)$ :  $place$  is along  $path$ .

- $isAlong(place, openspace)$ : For open spaces, since there is no beginning or end, places are only defined as being along an open space.
- $isAlong(place, corridor)$ : For corridors, the specific properties  $isAtBeginEdgeOfPath$ ,  $isAtEndEdgeOfPath$ ,  $isAtLeftOfPath$ ,  $isAtRightOfPath$  must be used. The choice of these properties is made with the arbitrary direction defined by positioning itself at its beginning and by traversing it towards its end.

*isBeside(place1, place2)*: *place1* is beside *place2*. Specified properties *isAtLeftOf* and *isAtRightOf* must be used to express the order of places. The choice of these properties is made by positioning themselves at the place and facing the path along the place.

*isInfrontOf(place1, place2)*: *place1* is in front of *place2*. This property does not need to be applied to all the places described. The more it is used, the more the verbalization of the itinerary will be easy. It is important to always define a place in front of an intersection to be able to determine if the guided human will have to go left or right in some cases. If there is no described place in front of an intersection, we can use a *emptyPlace* class that would inherit the *place* class.

The following axioms reduce the complexity of the description of the environment. The logical relations will therefore be solved by the ontology reasoner.

- $isAtLeftOf(place1, place2) \leftrightarrow isAtRightOf(place2, place1)$
- $isInfrontOf(place1, place2) \leftrightarrow isInfrontOf(place2, place1)$
- $isAlong(place, path) \wedge isIn(path, region) \rightarrow isIn(place, region)$

### 3 Computing routes

At this point, we have built an internal representation of the environment using the Semantic Spatial Representation (SSR). We illustrate how this representation can be used to compute the possible routes from one place to another. Even if the length of a route is taken into account in the choice, the complexity of the description is an important criterion (Morales et al., 2015). When someone asks his way, the guide will not necessarily try to give him the shortest way. His main goal is to make sure the person reach her goal. In the example of Figure 3, even if the red route is little longer than blue route, he will certainly propose it instead. Every intersection or change of direction is a risk for the guided person to make a mistake and thus get lost.

In this section, the goal is to provide multiple routes so that we can allow to choose the best route based on the person preferences. The possibility

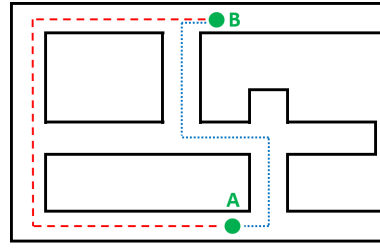


Figure 3: Comparison of two routes in terms of complexity and length. The blue route ( . . . ) is the shortest but is complex to explain while the red ( - - - ) is simpler although a bit longer.

of making this choice using the SSR will be presented in section 4. In order to reduce the complexity of the search, especially for large scale environments, we propose to work at two levels:

- **First** : Region level: considers only areas and passages such as doors, stairs or elevators.
- **Then** : Place level: provides complete routes description including paths and intersections within regions.

#### 3.1 Region level

In large-scale environments such as multi-storey buildings, routes computation can lead to combinatorial explosion. Exploration at the region level decreases this effect by conducting a first high-level exploration. In Figure 4 we can see that the exploration of paths of regions 4 and 5 is useless because these regions do not lead to the region containing the goal. This exploration uses only the regions and interface elements described in section 2.

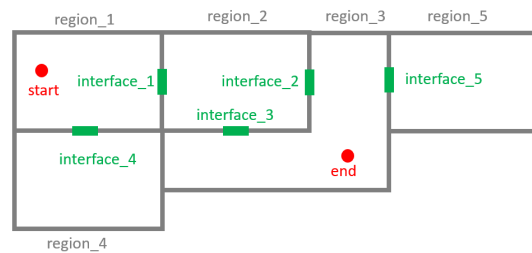


Figure 4: Representation of an environment at the regional level.

Each interface is connected to regions thanks to the *isIn* property. With this property, a route finding algorithm, based on the breadth-first search, makes possible to find the shortest routes connecting two regions by using the semantic knowledge.

By including the knowledge base exploration directly inside the search algorithms, it is not necessary to extract a topological graph with nodes and arcs. It is carried out within the search algorithm without preprocessing.

This algorithm applied to the example presented in Figure 4 gives the tree of Figure 5. The final routes found by the algorithm are :

- $region\_1 - interface\_1 - region\_2 - interface\_2 - region\_3$
- $region\_1 - interface\_1 - region\_2 - interface\_3 - region\_3$

Region 5 has never been explored and region 4 is not present in the final result. However, both solutions with interfaces 2 and 3 have been taken into account. This type of results makes possible to quickly eliminate unnecessary solutions and thus reduces the complexity for a more detailed search in a second time. This technique is similar to what is done for GNSS road navigation systems where the main roads are studied upstream of secondary roads with pyramidal (or hierarchical) route structure (Bovy and Stern, 1990).

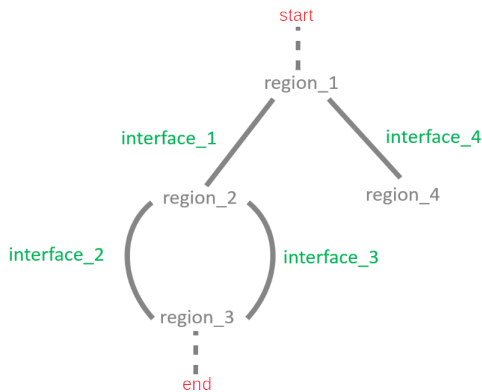


Figure 5: Exploration graph resulting from region-level search (sec.3.1) and the aggregation of start and end places (sec.3.2) .

### 3.2 Place level

Place-level search is based on the Region-level search results with the aggregation of start and end places, so the format changes from  $region - place - region - \dots - region$  to a  $place - region - place - \dots - place$ .

Place-level search works from one place to another through a single region. We have therefore divided the previous solutions to meet this constraint. This step aims to reduce complexity again.

Indeed, if several routes pass through the same region with the same places of departure and arrival, the inner route can be calculated once and for all. In our example, the division gives five sub-routes instead of six:

- $start - region\_1 - interface\_1$
- $interface\_1 - region\_2 - interface\_2$
- $interface\_2 - region\_3 - end$
- $interface\_1 - region\_2 - interface\_3$
- $interface\_3 - region\_3 - end$

The place-level algorithm aims to replace each sub-route region with a succession of paths and intersections. It works on the same principle as the previous search algorithm using the *isAlong* property instead of the *isIn* property. To improve performance, we use moving forward for the breadth-first search. It stops the exploration of the routes passing through a path already used in previous route calculation steps. In addition, it prevents loops.

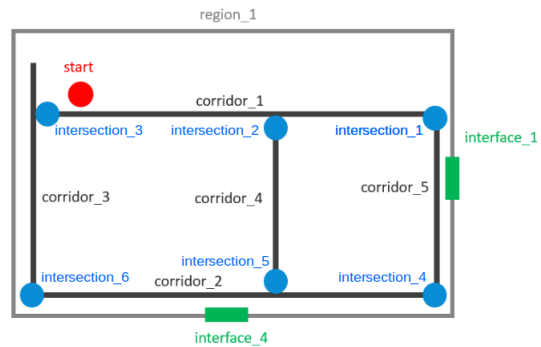


Figure 6: Representation of corridors and intersections in region 1 from the example 4

Taking the example of Figure 4 and focusing on region 1, we can solve the sub-route  $start - region\_1 - interface\_1$ . Region 1 is represented with its corridors and intersections in Figure 6. By applying the algorithm at the place level, we have the solution  $start - corridor\_1 - intersection\_1 - corridor\_5 - interface\_1$ . By doing the same for each sub-route, we can then re-compose the global routes and give the set detailed of routes from start to end.

## 4 Choosing the best route

Since the SRR is based on an ontology, we can have the meaning of each element of the environment and we can attach additional information to

them as features. Now that we have found several routes to the same goal, we want to select one based on different criteria. This selection of routes is independent of the previous section and a variety of cost functions can be implemented based on specific application needs. In the following subsection, we present an example of cost function using SSR and designed for robot guide to be deployed by the European project MuMMER (Foster et al., 2016).

#### 4.1 Example of cost function

As mentioned in (Morales et al., 2015), the complexity of the route to explain to a human, which is the number of steps in a route, is the most important criterion in choosing the route. A cost function taking into account only the complexity of a route  $R$  in the environmental context  $M$  would be  $f(R, M)$  with  $N$  being the number of steps of  $R$ .

$$f(R, M) = N \quad (1)$$

However, to find a good itinerary, it is important to take into account the preferences and capabilities of the guided person. An easy example is that we will never indicate a route with stairs to a person with reduced mobility. Using an ontology and therefore the possibility of describing the properties of elements of the environment, we add the property *hasCost* which associates an element with a criterion. Criteria rated  $\sigma_i$  are: saliency, accessibility, comfort, security, ease of explanation and speed. Other criteria could easily be added through the use of ontology according to the specific needs of the environment. All these criteria and their antonyms can be applied to each element  $n$ . The preferences of a person  $P$  are costs related to the criterion  $\sigma_i$  noted  $\varphi_{\sigma_i}$ . This represents the sensitivity of  $P$  to the  $\sigma_i$  criterion. The cost function becomes  $f(P, R, M)$  to take into account the preference of person  $P$ .

$$f(P, R, M) = N \times \prod_{n=0}^N [\prod_i (\sigma_{in} \times \varphi_{\sigma_i})] \quad (2)$$

Because we focused only on the complexity of the route explanation and the characteristics of the elements of the environment, in the presented cost function 2, the distances are not taken into account. This information could be added by working with a metric representation. Another possibility that can be explored is to add some of the

metric knowledge, such as the length of the corridor, into the semantic representation of the environment to preserve the working principle of a unique representation of the environment in this route description task.

## 5 Explanation generation

This section describes the third cognitive operation of (Denis, 1997) to generate a spatial discourse: the formulation of the procedure. As (Kopp et al., 2007), we define a route description as a set of route segments, each connecting two important points and explained in a chronological way. As (Tversky and Lee, 1999), we add to each route segment a triplet: orientation, action, and landmark to enable its explanation. The division into segments corresponds to all paths, with their entry and exit points, provided by our planning part. However, the semantic representation (SSR) used to plan the route is not directly usable to generate the formulation of the procedure. With the current representation, the orientation and action are too complex to extract (given that they depend on the direction by which the person arrives). It is however possible to interpret the semantic representation in relation to the estimated future position of the human. This interpretation is what we call the internal representation. This internal representation is composed of several sub-representations each representing a path of the global environment. Each segment of the route is represented independently of the others. For open space, we generate an ordered array of all locations *along* it. For the corridors, we generate four ordered arrays to represent the *left*, the *right*, the *beginedge* and the *endedge* of the corridors. These information can be found in the ontology with the properties *isAlong*, *isAtLeftOfPath*, and so on. To order the places in each array, we use the properties *isAtLeftOf* and *isAtRightOf* also present in the ontology. This internal representation can be displayed and gives Figure 7 for the corridor\_1 of region\_1 from the example 4. The *isInfrontOf* property is used to generate better placements.

Once we have an internal representation of each segment, we can determine the procedure that the user must perform. (Kopp et al., 2007) mention that an action, a reorientation, a progression or a positioning must be carried out at the end of each segment. The end of one

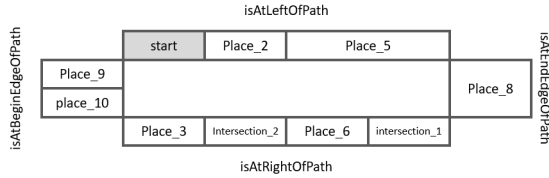


Figure 7: Internal representation of a the corridor\_1 of region\_1 from the example 6, extracted from the semantic representation.

segment being the beginning of the next, we choose to determine the actions at the beginning of each segment (which corresponds more to our internal representation). It allows to work on one path at a time. This rule is formalized as "choosing action  $A_i$  at place  $P_j$  will lead to place  $P_k$ " by (Mallot and Basten, 2009). This determination of actions can be made with our internal representation, as shown in Figure 8 where  $P_j$  is the gray place and  $P_k$  can be one of the other. The red information at the top gives the "turn right" action with  $P_k$  being *Place\_9* or *Place\_10*, "go in front of you" with  $P_k$  being *Place\_3* and "turn left" for the other places. The blue information on the sides gives the orientation of the sub-goal place  $P_k$  taking into account the previous reorientation. With this orientation information we can give an explanation of the form "take the corridor at your right" where the action is determined by the type of  $P_k$  and the side given by the orientation information. On the example of corridor\_1, to go from *start* to *Intersection\_1*, the full sentence will be *turn left then take the corridor at your right*. Moreover, by taking into account the orientation of the guided person after an action we allow to provide directions in the route perspective and so the guided person to perform an imaginary tour of the environment (Kopp et al., 2007).

By working segment by segment in the order given by the route search algorithm, we necessarily generate the explanations with a temporospatial ordering. This criterion is an important point in Allen's best practice in communicating route knowledge (Allen, 2000).

The latest critical information presented by (Tversky and Lee, 1999) is landmark and we have it in our representation. (Tversky and Lee, 1998) noted that more than 90% of the guiding spots on maps and verbal directions contained additional information, which corresponds also to



Figure 8: Resolution of directions and directions with an entry in the hallway by the gray square "start".

the results of (Denis, 1997) and the Allen's best practice (Allen, 2000). With our internal representation, we provide all the landmarks (corresponding to places because being defined as such) around which action must be taken and we can therefore refer to it to help the guided person. On the previous example, the sentence may be confusing because there are two corridors on the left. We are able to refer to *place\_8* which will be on the left or *place\_6* which will be on the right by projecting the future position of the human at *Intersection\_1*. With this new information, the situation can be disambiguated. The full sentence will become *turn left then take the corridor at your right straight after place\_6*.

The verbalization software based on the SSR and the principles described above was created based on a human-human study (Belhassein et al., 2017). Among the set of route description sentences, we have identified four types of explanatory components: those corresponding to the beginning of route description, to the end of a route, of progress in the route and the particular case of explanations with one step. These four types are only dependent of the position of the segment to be included in the global explanation procedure. For each type, we have identified various sub-types depending on the actions to be performed or the location of actions. For example, for the types of end-of-procedure sentences, we distinguish those where the end goal will be on the side, in front or at the current future position. In total, we have identified 15 sub-types. Each component of the explanation sentence has been classified into one of these sub-types. We want to be able to propose different ways to express the same things so the system does not have only one way to express the very same information. To represent similar sentences and to be able to generate sentences with variations, we have grouped sentences with close lexical structures. Each sentence is then represented with its variations as follows: ["you will"], ["see



”, ”find ”],[”it ”, ”/X ”], [”on ”], [”your ”, ”the ”], [”/D ”],[”side ”, ”when you walk ”, ”]. When using a sentence, the variations are randomly chosen with a uniform distribution. We can notice in the previous example the use of variables such that  $X$  which corresponds to the name of an element of the environment and  $D$  to a direction. We also used the  $Y$  variables for a reference points and  $DY$  for a reference point directions. If a sentence requires a variable that we have not been able to extract from our internal representation, then another sentence with the same meaning or another variation of the sentence that does not require the variable is chosen.

## 6 Applications

The SSR was first applied in an emulated mall to develop the algorithms <sup>1</sup>, but we also tested it in a real mall to study its applicability in a larger environment. Table 1 indicates the number of elements described in both environments. The number of places does not correspond only to the sum of the shops, interfaces and intersections because much more elements have been described, such as ATMs, restrooms or carts location.

	emulated	real
place	83	249
shop	19	135
interface	11	18
path intersection	10	52
path	11	42
region	5	4

Table 1: Number of elements described in the emulated and real environment.

Table 2 presents the CPU time needed for the routes computation and cost function algorithms for several cases, applied to the real environment representation. Even though specialized algorithms that work with a specific representation of the environment may be faster than ours, we can see here that they are acceptable in the context of a human robot interaction and especially in a route description task to both compute the path and verbalize it. Indeed, by providing semantic, topological and spatial knowledge within a single representation it can be used by several algorithms usually requiring different representations.

<sup>1</sup><https://cloud.laas.fr/index.php/s/Mvfty2xN9qymR2T>

We can also see the use of the regions to reduce the computation times with the two cases where three routes were found, one of the cases crossing one region and the other two.

Number of routes found	Number of regions crossed	Number of paths used	Path finding execution time (ms)
1	1	1	< 10
1	1	3	[20, 25]
3	1	9	[50, 55]
3	2	12	[30, 35]
16	2	75	[160, 170]
20	2	129	[180, 190]

Table 2: CPU time (min-max interval) time for computing routes in a big shopping mall description. Each row refers to a single request that can provide multiple routes to the goal.

To show the usability of the internal representation extracted from the SSR in the verbalization of the route, we have developed a software <sup>2</sup> that is able to verbalize the route found by our semantic planner. In examples of the sentences synthesized by this software (Table 3), we can see that for the same goal, it is possible to use different points of reference and to position them with respect to another element in the environment. All directions shown in the A and B examples take into account the future position of the guided human and provide indications from the perspective of the route.

Goal	Sentence
<b>Y</b>	You see there <b>Y</b> .
<b>Y</b>	It’s on the right side of <b>Z</b> .
<b>Y</b>	It’s on the left side of <b>X</b> .
<b>A</b>	Go through the door. Take the stairs at your left and turn left. Go almost at the very end of the corridor and, turn left at the door. After that you will see <b>A</b> on your right.
<b>B</b>	Go straight down this aisle. Then, walk to the very end of the corridor and it’s on the left there.

Table 3: Sentences generated by a software using the internal representation extracted from the SSR.

The applications presented previously have not

<sup>2</sup>[https://github.com/LAAS-HRI/route\\_verbalization](https://github.com/LAAS-HRI/route_verbalization)

only been tested as such, but have been integrated into a global robotic architecture and deploy in a mall center <sup>3</sup> as shown in figure 9. This integration shows that the results obtained by algorithms working with a single semantic representation of an environment are usable and are relevant in a more global task.



Figure 9: Robot describing a route to a human in a mall using the SSR and the associated algorithms. The sentence in green is the explanation of the route verbalized by the robot from the SSR representation: "just go down the corridor and then go almost at the very end of this corridor and it's on the left when you walk".

## 7 Conclusions

We have proposed an environment model that suitable to find routes, to choose one and to be able to verbalize it using a single representation. The key contribution of our work is the **semantic spatial representation (SSR)**, a formalization of how to describe an environment such as a large and complex public space mall using an ontology. We have also presented results about the use of our system by a robot that provides route guiding to humans in a shopping mall.

To benefit from our system, it could be interesting to integrate this representation and the corresponding algorithms to a dialog system (Papaioannou et al., 2018) in order to exploit more deeply its capacities. An interesting usage of this system already possible but not yet exploited because of the need of a dialog system, would be to use the guided person previous knowledge to choose a route and/or to generate an explanation ("If you know the place 2, from this one ..."). In the same vein, it would be possible to link it with a system such as Shared Visual Perspective Planner (Waldhart et al., 2018) to begin explaining the route from a visible point. This would reduce the length of the

<sup>3</sup><https://cloud.laas.fr/index.php/s/CJcPWmMU7TZGQJB>

explanations and thus ensure a better understanding of the itinerary for the guided person. Another improvement would be to use an ontology to ground the interaction (Lemaignan et al., 2012) as part of the route description task.

At this stage, only the topological representation has been integrated into the semantic representation. This is a good first step in working with a single representation that is easier to evolve and ensure consistency of knowledge. Future work would involve the integration of metric information, and thus geometric representation.

## Acknowledgments

This work has been supported by the European Unions Horizon 2020 research and innovation programme under grant agreement No. 688147 (MuMMER project).

## References

- Gary L. Allen. 2000. Principles and practices for communicating route knowledge. *Applied cognitive psychology*, 14(4):333359.
- Erik Andresen, David Haensel, Mohcine Chraïbi, and Armin Seyfried. 2016. Wayfinding and Cognitive Maps for Pedestrian Models. In *Traffic and Granular Flow '15*, pages 249–256. Springer International Publishing.
- Michael Beetz, Daniel Beler, Andrei Haidu, Mihai Pomarlan, and Asil Kaan Bozcuog. 2018. KnowRob 2.0 A 2nd Generation Knowledge Processing Framework for Cognition-enabled Robotic Agents. page 8.
- Kathleen Belhassen, Aurélie Clodic, H  l  ne Cochet, Marketta Niemel  , P  ivi Heikkil  , Hanna Lammi, and Antti Tammela. 2017. [Human-Human Guidance Study](#).
- Tim Berners-Lee, James Hendler, and Ora Lassila. 2001. [The Semantic Web](#). *Scientific American*, 284(5):34–43.
- P. H. Bovy and E. Stern. 1990. [Route Choice: Wayfinding in Transport Networks: Wayfinding in Transport Networks](#). Studies in Operational Regional Science. Springer Netherlands.
- Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk H  hnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. 1998. [The Museum Tour-Guide Robot RHINO](#). In *Autonome Mobile Systeme 1998, 14. Fachgespr  ch, Karlsruhe, 30. November - 1. Dezember 1998*, pages 245–254.
- Lili Cao and John Krumm. 2009. [From GPS Traces to a Routable Road Map](#). In *Proceedings of the 17th*

- ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09, pages 3–12, New York, NY, USA. ACM.
- Elizabeth R. Christil and William H. Warren. 2014. [From Cognitive Maps to Cognitive Graphs](#). *PLOS ONE*, 9(11):e112544.
- Aurelie Clodic, Sara Fleury, Rachid Alami, and al. 2006. [Rackham: An Interactive Robot-Guide](#). In *The 15th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2006, Hatfield, Herthfordshire, UK, 6-8 September, 2006*, pages 502–509.
- Allan M. Collins and M. Ross Quillian. 1969. [Retrieval time from semantic memory](#). *Journal of Verbal Learning and Verbal Behavior*, 8(2):240–247.
- Michel Denis. 1997. The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive*, 16:409–458.
- Mary Ellen Foster, Rachid Alami, Olli Gestra-nius, Oliver Lemon, Marketta Niemelä, Jean-Marc Odo-bez, and Amit Kumar Pandey. 2016. [The MuM-MER Project: Engaging Human-Robot Interaction in Real-World Public Spaces](#). In *Social Robotics. 8th International Conference, ICSR 2016, Kansas City, MO, USA, November 1-3, 2016 Proceedings*, Lecture Notes in Computer Science, pages 753–763.
- Sachithra Hemachandra, Matthew R. Walter, Stefanie Tellex, and Seth Teller. 2014. [Learning spatial-semantic representations from natural language descriptions and scene classifications](#). In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2623–2630, Hong Kong, China. IEEE.
- Stefan Kopp, Paul A. Tepper, Kimberley Ferriman, Kristina Striegnitz, and Justine Cassell. 2007. [Trading Spaces: How Humans and Humanoids Use Speech and Gesture to Give Directions](#). In Toyoaki Nishida, editor, *Wiley Series in Agent Technology*, pages 133–160. John Wiley & Sons, Ltd, Chichester, UK.
- B. Kuipers, J. Modayil, P. Beeson, M. MacMahon, and F. Savelli. 2004. [Local metrical and global topological maps in the hybrid spatial semantic hierarchy](#). In *IEEE ICRA, 2004. Proceedings. ICRA '04. 2004*, volume 5, pages 4845–4851 Vol.5.
- Benjamin Kuipers. 2000. [The Spatial Semantic Hierarchy](#). *Artificial Intelligence*, 119(1):191–233.
- Sylvain Lefebvre and Samuel Hornus. 2003. [Automatic Cell-and-portal Decomposition](#). report, INRIA.
- Séverin Lemaignan, Raquel Ros, Emrah Akin Sisbot, Rachid Alami, and Michael Beetz. 2012. [Grounding the Interaction: Anchoring Situated Discourse in Everyday Human-Robot Interaction](#). *International Journal of Social Robotics*, 4(2):pp.181–199. 20 pages.
- Bing Liu. 1997. [Route finding by using knowledge about the road network](#). *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 27(4):436–448.
- Bernhard Lorenz, Hans Jrgen Ohlbach, and Edgar-Philipp Stoffel. 2006. [A Hybrid Spatial Model for Representing Indoor Environments](#). In *Web and Wireless Geographical Information Systems*, Lecture Notes in Computer Science, pages 102–112. Springer Berlin Heidelberg.
- Hanspeter A. Mallot and Kai Basten. 2009. [Embodied spatial cognition: Biological and artificial systems](#). *Image and Vision Computing*, 27(11):1658–1670.
- Yoichi Morales, Satoru Satake, Takayuki Kanda, and Norihiro Hagita. 2015. [Building a Model of the Environment from a Route Perspective for Human-Robot Interaction](#). *International Journal of Social Robotics*, 7(2):165–181.
- Luis Yoichi Morales Saiki, Satoru Satake, Takayuki Kanda, and Norihiro Hagita. 2011. [Modeling Environments from a Route Perspective](#). In *Proceedings of the 6th International Conference on Human-robot Interaction, HRI '11*, pages 441–448, New York, NY, USA. ACM. Event-place: Lausanne, Switzerland.
- Yusuke Okuno, Takayuki Kanda, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. [Providing route directions: design of robot's utterance, gesture, and timing](#). In *HRI, 2009 4th ACM/IEEE International Conference on*, pages 53–60. IEEE.
- Ioannis Papaioannou, Christian Dondrup, and Oliver Lemon. 2018. [Human-robot interaction requires more than slot filling - multi-threaded dialogue for collaborative tasks and social conversation](#). In *FAIM/ISCA Workshop on Artificial Intelligence for Multimodal Human Robot Interaction*. ISCA.
- Satoru Satake, Kotaro Hayashi, Keita Nakatani, and Takayuki Kanda. 2015a. [Field trial of an information-providing robot in a shopping mall](#). In *2015 IEEE/RSJ IROS*, pages 1832–1839.
- Satoru Satake, Keita Nakatani, Kotaro Hayashi, Takyuki Kanda, and Michita Imai. 2015b. [What should we know to develop an information robot?](#) *PeerJ Computer Science*, 1:e8.
- Roland Siegwart, Kai Oliver Arras, Samir Bouabdallah, Daniel Burnier, Gilles Froidevaux, Xavier Grep-pin, Björn Jensen, Antoine Lorotte, Laetitia Mayor, Mathieu Meisser, Roland Philippsen, Ralph Pignet, Guy Ramel, Gregoire Terrien, and Nicola Tomatis. 2003. [Robox at Expo.02: A large-scale installation of personal robots](#). *Robotics and Autonomous Systems*, 42(3-4):203–222.

- Kamil Skarzynski, Marcin Stepniak, Waldemar Bartyna, and Stanislaw Ambroszkiewicz. 2017. **SO-MRS: a multi-robot system architecture based on the SOA paradigm and ontology**. *arXiv:1709.03300 [cs]*. ArXiv: 1709.03300.
- Sebastian Thrun. 2008. **Simultaneous Localization and Mapping**. In Margaret E. Jefferies and Wai-Kiang Yeap, editors, *Robotics and Cognitive Approaches to Spatial Mapping*, Springer Tracts in Advanced Robotics, pages 13–41. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Rudolph Triebel, Kai Arras, Rachid Alami, and al. 2016. **SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports**. In David S. Wettergreen and Timothy D. Barfoot, editors, *Field and Service Robotics: Results of the 10th International Conference*, Springer Tracts in Advanced Robotics, pages 607–622. Springer International Publishing, Cham.
- Barbara Tversky and Paul U. Lee. 1998. **How Space Structures Language**. In Christian Freksa, Christopher Habel, and Karl F. Wender, editors, *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, Lecture Notes in Computer Science, pages 157–175. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Barbara Tversky and Paul U. Lee. 1999. Pictorial and Verbal Tools for Conveying Routes. In *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, Lecture Notes in Computer Science, pages 51–64. Springer Berlin Heidelberg.
- Jules Waldhart, Aurlie Clodic, and Rachid Alami. 2018. Planning Human and Robot Placements for Shared Visual Perspective. page 10.
- H. Zender, O. Martnez Mozos, P. Jensfelt, G. J. M. Kruijff, and W. Burgard. 2008. **Conceptual spatial representations for indoor mobile robots**. *Robotics and Autonomous Systems*, 56(6):493–502.