Verbalization: Narration of Autonomous Robot Experience

Stephanie Rosenthal
Software Engineering Institute
srosenthal@sei.cmu.edu

Sai P. Selvaraj
Robotics Institute
spandise@andrew.cmu.edu

Manuela Veloso
Computer Science Department
veloso@cs.cmu.edu

Carnegie Mellon University, Pittsburgh USA

Abstract

Autonomous mobile robots navigate in our spaces by planning and executing routes to destinations. When a mobile robot appears at a location, there is no clear way to understand what navigational path the robot planned and experienced just by looking at it. In this work, we address the generation of narrations of autonomous mobile robot navigation experiences. We contribute the concept of *verbalization* as a parallel to the well-studied concept of visualization. Through verbalizations, robots can describe through language what they experience, in particular in their paths. For every executed path, we consider many possible verbalizations that could be generated. We introduce the *verbalization space* that covers the variability of utterances that the robot may use to narrate its experience to different humans. We present an algorithm for segmenting a path and mapping each segment to an utterance, as a function of the desired point in the verbalization space, and demonstrate its application using our mobile service robot moving in our buildings. We believe our verbalization space and algorithm are applicable to different narrative aspects for many mobile robots, including autonomous cars.

1 Introduction

Service robots can autonomously generate and execute plans to successfully perform tasks for humans, appropriately handling the uncertainty of their surroundings. With mobile robots performing more autonomous behaviors without human intervention, humans in the environment may wonder what exactly the robot was perceiving, predicting, planning, and doing. Robotics researchers have developed logging approaches to enable the recording of the robot experience. For debugging purposes, such developers must dig through the accumulated robot logs to find out about the robot experience in great detail. In addition to researchers, an office worker may want the robot to identify why it was late in completing its task. And a person accompanying the robot may want the robot to summarize its speed and distance traveled. To the authors' knowledge, there are no robots that currently narrate in plain English their planned and executed experiences through a translation of sensor data and plans into natural language. In this work, we introduce *verbalization* as the process of converting or narrating robot experiences via natural language. A robot that verbalizes its experiences could help each of the above example users resolve questions they have about autonomous robot behavior.

Different humans interacting with autonomous robots, as exemplified above, are interested in different specific information, for specific parts of the robot's experience, and at different levels of detail. A one-size-fits-all verbalization will not satisfy all users. We contribute the concept of the *verbalization space* to represent ways in which verbalizations may vary for different reasons, including user preferences and needs. We define our verbalization space across three orthogonal parameters that prior research has indicated per-user needs or preferences over [Dey, 2009; Bohus *et al.*, 2014; Thomason *et al.*, 2015]. The first parameter, *abstraction*, varies the vocabulary and concepts used in the narrative from concrete robot concepts, such as distances, speed, and time to abstract concepts, such as hallways, rooms, landmarks. Second, *specificity* varies the total number of concepts or words used in the summaries, allowing the robot to generate single-sentence general, or multi-sentence detailed, narratives. Finally, *locality* varies the particular parts of the experience that the narration focuses on, from the global path to a local region or landmark of interest. Our verbalization space is general and can be extended to many other parameters.

We first formalize the concept of verbalizing experiences, as well as each of the parameters of our verbalization space with a focus on navigation tasks. We contribute our algorithm for generating narratives using the three verbalization space parameters, and we provide examples of how to combine these parameters. Our algorithm can be adapted to use other natural language generation techniques or verbalization space parameters. Finally, we demonstrate the use of our verbalization space to narrate our mobile robot's experiences through our building, and validate that it generates narratives of different abstraction, specificity, and locality.

2 Related Work

Prior work in automatically generating explanations or summaries of planned behavior can be roughly divided into three categories: 1) intelligibility or explanation of machine learning algorithms, 2) summarizing perceived behavior, and 3) generating directions for humans to follow.

As machine learning gains popularity in many different applications, much human-computer interaction research has focused on ways machine learning applications can *intelligibly* explain their reasoning algorithms to users (e.g., for context-aware systems [Dey, 2009]). HCI intelligibility studies have focused on ways that users can query applications for information or explanations (e.g., [Lim *et al.*, 2009]) as well as how those explanations can affect users' mental models of how the applications work (e.g., [Kulesza *et al.*, 2012; 2013]). The studies find that explanations increase trust of machine learning applications [Bussone *et al.*, 2015] as well as improve users' mental models. Due to the success of intelligibility across many applications, intelligibility toolkits have been implemented for consistency of explanation across different machine learning algorithms [Lim and Dey, 2010]. While prior work shows that varying the focus of explanations is important and useful to users, no one implements it.

Another growing area of research is in summarizing or generating narratives of perceived behavior. For example, RoboCup soccer commentators aim to use the input of simulated RoboCup games [Voelz *et al.*, 1999] or live RoboCup games [Veloso *et al.*, 2008] to generate realtime summaries of the actions in the games. Activity recognition algorithms and natural language generation have also been used to produce annotated accounts of wartime exercises [Luotsinen *et al.*, 2007], video conferencing sessions [Yengui and Mahmoud, 2009], and sports games [Allen *et al.*, 2010]. While some work generates a variety of summaries to maintain human interest (e.g., [Veloso *et al.*, 2008]), the work does not vary the length or depth of summaries as we do.

Finally, and perhaps most closely related to our work, GPS applications (e.g., [Belvin *et al.*, 2001]) and robot applications (e.g., [Kirby *et al.*, 2005; Bohus *et al.*, 2014; Thomason *et al.*, 2015]) are automatically generating navigation instructions and dialog for people to follow and understand. In the prior work, a path is converted into language and ideally presented in an easy-to-understand yet accurate way for the person to follow it seamlessly every time. While these navigation directions do not vary in the language used, recently [Bohus *et al.*, 2014] found that navigation directions should 1) provide differing levels of specificity at different locations in the route and 2) use abstract landmarks in addition to more concrete details. Similarly, prior work on human direction givers shows that humans do not generate the same directions for every person [MacFadden *et al.*, 2003].

We note that none of the prior work focuses on summarizing both perception and plans of a robot or other autonomous vehicle. And while the prior work extensively documents the need for parameterized summaries, none of the prior work, to our knowledge, measures those parameters and contributes an algorithm for actually varying them. In this work, we first contribute verbalization as a method of summarizing what robots actually experience. Based on the findings from prior work as well as the needs of our robots' users, we then propose and formalize our verbalization space that represents the variability in narratives, and we provide an algorithm for generating variable verbalizations of route plans.

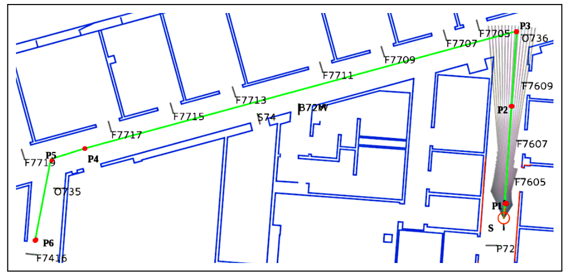

Figure 1: Robot route plan (green lines), nodes $\{S,P_1,...P_6\}$, Starting node S, and finish node P_6.

3 Route Verbalization

We define *verbalization* as the process by which an autonomous robot converts its own experience into language. In this work, we consider mobile navigation experience in the physical world, and verbalize what the robot experienced while traversing its route. We define *route verbalization* as the process by which an autonomous robot converts its own route experience into language. A robot can generate route verbalizations mentioning the planned route that will be traversed or the route that has been traversed (*i.e.,* a narrative in the future tense is equivalent to GPS driving directions, while a narrative of the past traversed route describes the actual experience). At this time, we do not distinguish between the future and past tenses, exemplifying the applicability across language generation domains.

We first define simple route verbalizations over common robot map and route representations. Then, we describe our annotations to the map and route to accommodate the variation in verbalization that humans require.

3.1 Robot Map and Route Plan

We define an indoor mobile robot's map $M = \langle P, E \rangle$ as set of points $p = (x, y, b, z) \in P$ representing unique locations (x, y) in our buildings b for each floor z and edges $e = \langle p_1, p_2, d, t \rangle \in E$ that connect that connect two points taking time t to traverse distance d.

The points on the map are annotated with semantic *landmarks* represented as *room numbers* (e.g., 7412, 3201) and *room type* (office, kitchen, bathroom, elevator, stairs, other). Points could be annotated with additional information, including the occupants of the office or the names of laboratory spaces (e.g., as in [Rosenthal *et al.*, 2010]). We also maintain lists of *corridors* and *bridges* as points that reside within them (e.g., "7400 corridor" contains office 7401, office 7402, office 7404, etc. and the "7th floor bridge" contains other 71, other 72, etc.). Some points may not appear in any corridor or bridge list if they are in open areas, and some points may reside in two hallways if they occur at hall intersections.

Using our map, our route planner produces plans as trajectories through the environment composed of:

- a starting point S,
- a finish point F,
- an ordered list of intermediate waypoints $W \subset P$, and

Specificity, S		Abstraction, A			
		Level 1	Level 2	Level 3	Level 4
	General Picture	Start and finish point of the complete route	Total distance and time taken for the complete route	Total distance and time taken for the complete route	Starting and ending landmark of complete route
	Summary	Start and finish point for subroute on each floor of each building	Total distance and time taken for subroute on each floor of each building	Total distance and angles for subroute on each floor of each building	Starting and ending landmark for subroute on each floor of each building
	Detailed Narrative	Start and finish points of complete route plus time taken for each edge of route	Angle turned at each point plus the total distance and time taken for each edge of route	Turn direction at each point plus total distance for each edge of route	All landmarks encountered on the route

Table 1: Narrated information depends on preferred Verbalization Space parameters. Information for Abstraction A and Specificity S are shown assuming Locality L is Global. For a different Locality, a subset of the route is generated, and the information provided is computed in terms of the subset.

- a subset of straight line edges in E that connect S to F through W.

Our planner labels waypoints as *turning points* representing the only places the robot turns after traversing straight edges. Figure 1 shows a route plan, the starting point S, and finish point $F = P_6$, as the destination of a task requested by a user. The figure shows turning points $W = \{P_1, P_2, P_3, P_4, P_5\}$, connected by straight line edges (as pictured in green).

3.2 Simple Route Verbalization

Using the map and route plan described above, a simple route verbalization algorithm could interleave turn angles at each point p and distances traversed for each edge e between waypoints. For the route depicted in Figure 1, this simple route verbalization algorithm would produce:

> *I went straight for 8.5 meters and turned left, then straight for 24.9 meters and turned left, then straight for 3.62 meters to reach the destination.*

While this verbalization successfully describes the robot's route, different people in the environment may be expecting more or different information to be provided. For example, we as robotics researchers could be interested in the exact (x, y, b, z) coordinates of the points where the robot turns. Other people in the environment may find landmarks such as room numbers to be useful. We next describe the use of our semantic annotations within our verbalization space.

4 Verbalization Space

We represent the variations in possible narratives of the same route as the *verbalization space*. Each region of the verbalization space represents a different way to generate text to describe the route plan. A user may specify their personalized preferences for verbalization within this space, or the preferences may be inferred from some other source. Our verbalization space contains three orthogonal parameters – abstraction, locality, and specificity – that are well-documented as personal preferences in the literature (e.g., [Dey, 2009; Bohus *et al.*, 2014; Thomason *et al.*, 2015]). Our verbalization space is general and could be extended to include more parameters as needed.

4.1 Verbalization Space Definitions

Table 1 details the way we instantiate verbalizations for specified parameters $(a, l, s) \in (A, L, S)$.

Abstraction A: Our abstraction parameter represents the vocabulary or corpus used in the text generation. In the most concrete form (Level 1), we generate text in terms of the robot's world representation, directly using points (x, y, b, z) from the route plan. Our Level 2 derives turn angles and uses expected or actual traversal time and distances from the points and edges in the plan. Level 3 abstracts the angles and distances into right/left turns and straight segments. And finally, in the highest level of abstraction, Level 4 contains the semantic annotations described above.

Locality L: Locality describes the segment(s) of the route the user is interested in. In the most general case, the user is interested in the route through the entire Global Environment including all buildings and floors. However, an office occupant may only be interested in a particular predefined Region of the route composed of multiple points in the maps (e.g., we limit our regions by building b or building floor b, z). Finally, the occupant may specify a single particular point or landmark for the robot to summarize its route around (e.g., a constant distance around the 8th floor kitchen or Office 4002).

Specificity S: Specificity indicates the number of concepts or details to discuss in the text: the General Picture, the Summary, and the Detailed Narrative. The General Picture contains the most general description of the robot's route, namely the start and finish points (or landmarks), the total distance covered, and/or the time taken (see Table 1). Our Summaries contain this same information for the subroute on each floor of each building. The Detailed Narrative contains a description of each edge of the robot's route.

Next we describe how these verbalization space parameters are used to generate verbalization text.

Algorithm 1 Variable Verbalization Algorithm

Input: *route, verb_pref, map* **Output:** *narrative*

```
   //The verbalization space preferences
1: (a, l, s) ← verb_pref
   //Choose which abstraction vocabulary to use
2: corpus ← ChooseAbstractionCorpus(a)
   //Annotate the route with relevant map landmarks
3: annotated_route ← AnnotateRoute(route, map, a)
   //Subset the route based on preferred locality
4: subset_route ← SubsetRoute(annotated_route, l)
   //Divide the route into segments, one per utterance
5: route_segs ← SegmentRoute(subset_route, s)
   //Generate utterances for each segment
6: utterances ← NarrateRoute(route_segs, corpus, a, l, s)
   //Combine utterances into full narrative
7: narrative ← FormSentences(utterances)
```

4.2 Variable Verbalization Algorithm

The Variable Verbalization (VV) algorithm pseudocode is presented in Algorithm 1. The algorithm directly translates the robot's route plan into plain English given the map and the incorporated annotations described above. It takes as input a *route*, a verbalization space preference *verb_pref* = $(a, l, s) \in (A, L, S)$, and a *map* of the environment with locations labeled as above. It starts by choosing what corpus (Level 1-4) to use when generating utterances depending on abstraction preference a (Line 2). Then, the VV algorithm annotates the given route by labeling each point with landmarks and corridor/bridge names using the map (Line 3).

Once the route is annotated with relevant locations, the algorithm extracts the subset of the route that is designated as relevant by the locality preference l (Line 4). We subset Regions by building and floor and Landmarks by a threshold distance around a given point. Both of these subset types can be directly computed from our point representation - Regions using b, z and Landmarks using a distance function around x, y for the given building/floor. The output of this step is another annotated route that is a copy of the route if l=Global Environment. Otherwise, the output is a subset of the route with a new start and finish point.

Using the *subset_route*, the VV algorithm then computes route segments to narrate with respect to the specificity preference s (Line 5). If the specificity preference is a General Picture, our algorithm computes the required abstraction information for a single route segment from S to F. For Summaries, it computes one route segment for each floor of each building and then computes the relevant abstraction information for those segments. In Detailed Narratives, all edges are included in the narrative.

The Algorithm then translates the route segments from Line 5 into plain English using the corpus vocabulary from the annotated map and template sentences (Line 6, examples described next). Finally, after the sentences have been generated for each route segment, the VV algorithm stitches them together (Line 7). The final narrative is returned as the output of the function.

In the next section, we describe our implementation of our algorithm on our mobile robot and its routes.

5 Mobile Robot Route Verbalizations

Our mobile service robot plans and executes tasks autonomously in our buildings [Biswas and Veloso, 2013; 2014], such as accompanying visitors to their meetings and carrying objects to offices [Veloso *et al.*, 2015]. It regularly interacts with humans in the environment through dialog and symbiotic interactions to ask for help [Rosenthal *et al.*, 2010; Perera *et al.*, 2015; Perera and Veloso, 2015]. We found many different people in our environment are interested in what our robot is doing and experiencing as it acts. We as researchers tend to be interested in high specificity, detailed narratives about the global environment. Other people may be interested in narratives about their own office locations at a general picture level. The Variable Verbalization algorithm is implemented on our robot and allows each person to receive a personalized narrative based on their priorities and interests.

We first describe our annotated map and corpus for verbalizations that are input into our Variable Verbalization algorithm. Then, we describe two narratives based on different verbalization space preferences for the same route. Finally, we test our algorithm on different routes through our building to demonstrate how the number of words and numbers changes with each instantiation of our verbalization space.

5.1 Robot Map and Language Corpus

Our robot's environment includes three buildings connected by bridges. Each floor of each building has its own coordinate system. The individual floor maps are linked to each other via the elevators and bridges, so that the robot can use multiple floors while planning and executing. The set of all floors and all buildings is defined as our map M. Our map contains points p representing any arbitrary location on the map. Points can be labeled as landmarks representing specific room numbers and room types including office, lab, kitchen, bathroom, elevator, stairs, printers, and other. We also maintain lists of corridors and bridges as outlined above. Given any two points, start S and finish F, our route planner computes a set of edges and waypoints to travel from S to F.

Our corpus of landmarks on the map (exerpt below) is used for Level 4 of our Abstraction parameter. Our other corpora for our other levels of abstraction are much smaller and include (x, y) "points", "angle" degrees, distance in "meters", "left turns", "right turns", and "u-turns".

$$\begin{cases} \ldots & \ldots \\ \text{Office-3201} & (x, y, Gates, 3^{rd}\ floor) \\ \text{Bathroom-3} & (x, y, Gates, 3^{rd}\ floor) \\ \text{Stairs-34} & (x, y, Gates, 3^{rd}\ floor) \\ \text{Kitchen-71} & (x, y, Gates, 7^{th}\ floor) \\ \text{Office-7401} & (x, y, Gates, 7^{th}\ floor) \\ \text{Office-7412} & (x, y, Gates, 7^{th}\ floor) \\ \ldots & \ldots \end{cases}$$

5.2 Route Experience Variable Verbalization

Using our map, our mobile robot plans routes between points in our building. Figure 2 Top shows one example route (in green) from the 3rd Floor Office 3201 to the 7th Floor Office 7416 in our Gates building. We have labeled in black our annotations over the map including the corridors, the elevators, a bridge, and a kitchen. Figure 2 Bottom shows a

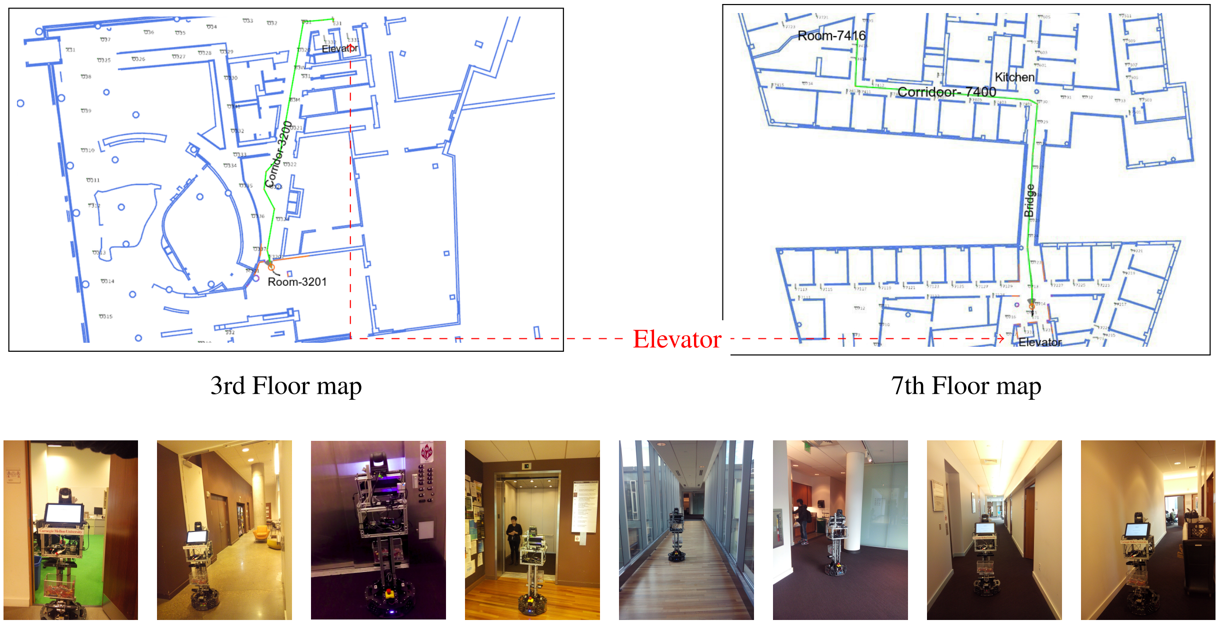

Figure 2: Top: Example of our mobile robot's multi-floor plan in our building (blue walls, green route, red connects elevator between floors). Bottom: Images of our robot navigating the route. The robot (1) starts at Office 3201, (2) travels down the 3200 corridor, and turns right to (3) reach the elevator. Once it (4) reaches the 7th floor, it (5) travels straight across the bridge, (6) turns left at the kitchen, (7) travels down the 7400 corridor, and then (8) makes its first right to Office 7416.

visual depiction of the robot traveling along this route. We demonstrate two variations of verbalizations for the route.

Example 1: Long, Detailed Verbalization

With our map and corpus, we consider the preference:

(Level 4, Global Environment, Detailed Narrative)

that represents a researcher in our lab who wants a detailed description of what happens on each edge of the robot's route. We will review our algorithm's analysis of the route plan to generate a verbalization fitting this preference.

Choose Abstraction Corpus: Because the abstraction level preference is Level 4, the VV algorithm chooses the large corpus of room numbers, room types, and corridors and bridges for its language model.

Annotate Route: Next, the input route is annotated with these landmarks from the corpus. In this case, the VV algorithm labels starting point Office-3201; the points leading to the elevator are Corridor-3200; the elevator on the 3rd floor is labeled Elevator-31 and similarly the 7th floor is labeled Elevator-71; points on the bridge are Bridge-7; the Kitchen-71 is labeled; the hallway points are labeled Corridor-7400; and finally the finish point is Office-7416.

Subset Route: The researcher is interested in the Global Environment Locality, and thus the route is not subsetted.

Segment Route: The researcher would like s =Detailed Narrative. Our algorithm merges all same-labels, resulting in seven route segments. We write segments in terms of their meaning here because there are too many points to enumerate; the robot maintains the list of points on the route.

{s1: Office-3201, s2: Corridor-3200, s3: Elevator, s4: Bridge-7, s5: Kitchen-71, s6: Corridor-7400, s7: Office-7416}

Narrate Route: Our algorithm's ability to narrate a route depends on filling in templates matching different route segments. We manually created the following templates for Level 4 abstractions. We note next to the D whether the type of landmark is specific (e.g., the template must be filled in by a corridor, bridge, etc.), and we note with a slash that the choice of verb is random to prevent repetition by replacing the verbs with a synonym (e.g., [Veloso *et al.*, 2008]). We have similar templates for other abstraction levels that include distances and time to complete the route segments.

- "$[\text{I}]_N$ $[\text{visited/passed}]_V$ the $[\_\_\_]_{D:room}$"
- "$[\text{I}]_N$ $[\text{took}]_V$ the elevator and went to the $[\_\_\_]_{D:floor}$"
- "$[\text{I}]_N$ $[\text{went through/took}]_V$ the $[\_\_\_]_{D:corridor/bridge}$"
- "$[\text{I}]_N$ $[\text{started from}]_V$ the $[\_\_\_]_{D:start}$"
- "$[\text{I}]_N$ $[\text{reached}]_V$ $[\_\_\_]_{D:finish}$"

Using the templates, the VV Algorithm generates utterances for each of the segments.

{
s1: "I started from Office 3201",
s2: "I went through the 3200 corridor",
s3: "I took the elevator to the seventh floor",
s4: "I took the 7th floor bridge",
s5: "I passed the kitchen",
s6: "I went through the 7400 corridor",
s7: "I reached Office 7416",

Form Sentences: Finally, the algorithm combines the sentences with "then"s (more complex concatenation could be used):

I started from office 3201, then I went through the 3200 corridor, then I took the elevator and went to the seventh floor, then I took the 7th floor bridge, then I passed the kitchen, then I went through the 7400 corridor, then I reached office 7416.

Example 2: Short Overview Verbalization
To contrast the long detailed landmark-based narrative, a short verbalization can be achieved with preference

(Level 2, Gates 7th Floor Region, General Picture)

Here, a person accompanying the robot wants to know how far they traveled only on the 7th floor. The VV algorithm first annotates our entire route with abstraction Level 2, adding distances to the edges in the route between each pair of points. Since the required locality is Region, the algorithm subsets the route containing only the required Gates 7th floor points. As the specificity is General Picture, a single route segment is generated as the combination of all edges from the new 7th floor start node S to the finish node F. The route is annotated with the total distance and time taken for the route. Next the algorithm narrates the route using the template "$[I]_N$ [traveled] [x] meters in [t] seconds on the $[\_\_]_{D:floor}$". Finally these utterances could be combined (not necessary here) to form the final narrative:
I traveled 56.18 meters and took 75 seconds on the 7th floor.

5.3 Validation

Given the well-documented need for verbalizations, we focus our experiment on whether we succeed at varying our verbalizations based on those needs. We randomly generated 12 multi-floor routes in our Gates building and 12 single-floor routes, ran the VV algorithm over the route plans, and analyzed the content of the 36×24 verbalizations that were generated.

Figure 3 shows the average number of words for two of our parameters: abstraction and specificity. There are many more words in Detailed Narratives (55-104 words) compared to Summaries (14-21) or General Pictures (10-18). We note that the number of words is nearly the same for Summaries and General Pictures. Because our VV implementation creates one phrase per floor of the building for Summaries, it generates the same narrative as the General Picture for single-floor navigation routes. Given that half of our routes are single-floor, the average number of words for Summaries is similar to that of General Picture rather than Detailed Narratives.

Additionally, there are more words generated for Summary/General Picture Level 4 Abstraction than Level 3 or 2. This is due to the landmark descriptions that are more verbose than the time and distances reported. In contrast, for Abstraction Level 4, there are no numbers in most of our narratives as the landmarks are entirely made up of words (Figure 4). The exception is Level 4 Abstractions with Detailed Narratives, which do include office numbers.

The addition of the locality parameter reduces the overall number of words and numbers but shows the same patterns. As the narratives become more focused around a region and then a landmark, there are fewer route segments to describe. We conclude that overall we do successfully vary narratives within our verbalization space.

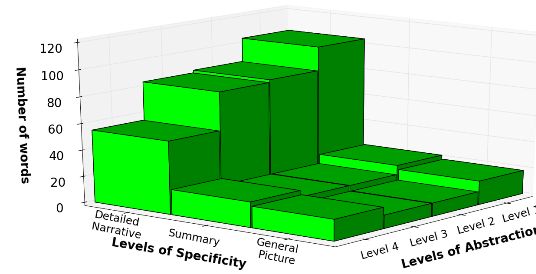

Figure 3: Average number of words generated.

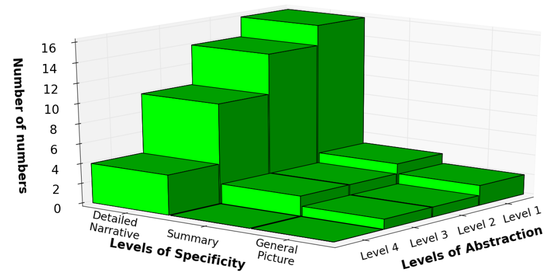

Figure 4: Average number of numbers generated.

6 Conclusion

It is hard, if not impossible, for humans to understand the experience of an autonomous mobile robot. In this paper, we have contributed a novel approach to capture verbalization by a robot as a way for the robot to narrate its experience in natural language. Our mobile robot translates its route experiences in into verbalization utterances. We contribute the verbalization space as a formalization of multiple levels of detail in which narrations can be generated. We introduce different axes of the space to represent different dimensions of verbalization, namely abstraction, locality, and specificity, though the space can be extended.

The approach we present aims at being applicable beyond mobile robots to other planning algorithms, allowing language to be adjusted to the desired levels of detail. For autonomous vehicles, we can imagine using a new map and semantic landmark labels with our same verbalization space and the same verbalization algorithm to produce narrations of driven routes. Autonomous vehicles would reason over points in GPS space, and use landmarks such as buildings, roads, and street signs to create a variety of narrations. Other intelligible machine learning applications could also produce new formalisms for the verbalization space to produce variable narrations.

We demonstrate the use of verbalizations on our mobile service robot. We present two examples of narrations corresponding to different points in the verbalization space for one multi-floor route through our building environment. Then, we validate on 24 routes that a variety of narrations that can be generated from any single plan. Future work will focus on studying techniques for the personalization of verbalization preferences among our building occupants.

Acknowledgments

Copyright 2016 Carnegie Mellon University. This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT. [Distribution Statement A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution. DM-0003247

References

[Allen *et al.*, 2010] Nicholas D Allen, John R Templon, Patrick Summerhays McNally, Larry Birnbaum, and Kristian J Hammond. Statsmonkey: A data-driven sports narrative writer. In *AAAI Fall Symposium: Computational Models of Narrative*, 2010.

[Belvin *et al.*, 2001] Robert Belvin, Ron Burns, and Cheryl Hein. Development of the hrl route navigation dialogue system. In *Proceedings of the first international conference on Human language technology research*, pages 1–5. Association for Computational Linguistics, 2001.

[Biswas and Veloso, 2013] Joydeep Biswas and Manuela Veloso. Localization and Navigation of the CoBots over Long-term Deployments. *International Journal of Robotics Research*, 32(14):1679–1694, 2013.

[Biswas and Veloso, 2014] Joydeep Biswas and Manuela Veloso. Episodic Non-Markov Localization: Reasoning about Short-Term and Long-Term Features. In *Proceedings of the International Conference on Robotics and Automation*, 2014.

[Bohus *et al.*, 2014] Dan Bohus, Chit W Saw, and Eric Horvitz. Directions robot: In-the-wild experiences and lessons learned. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems*, pages 637–644, 2014.

[Bussone *et al.*, 2015] Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. The role of explanations on trust and reliance in clinical decision support systems. *IEEE International Conference on Healthcare Informatics*, 2015.

[Dey, 2009] Anind K Dey. Explanations in context-aware systems. In *ExaCt*, pages 84–93, 2009.

[Kirby *et al.*, 2005] Rachel Kirby, Frank Broz, Jodi Forlizzi, Marek Piotr Michalowski, Anne Mundell, Stephanie Rosenthal, Brennan Peter Sellner, Reid Simmons , Kevin Snipes, Alan Schultz, and Jue Wang. Designing robots for long-term social interaction. In *Proceedings of the International Conference on Intelligent Robots and Systems*, pages 2199 – 2204, 2005.

[Kulesza *et al.*, 2012] Todd Kulesza, Simone Stumpf, Margaret Burnett, and Irwin Kwan. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1–10. ACM, 2012.

[Kulesza *et al.*, 2013] Todd Kulesza, Simone Stumpf, Margaret Burnett, Songping Yang, Irwin Kwan, and Weng-Keen Wong. Too much, too little, or just right? ways explanations impact end users' mental models. In *Visual Languages and Human-Centric Computing (VL/HCC), 2013 IEEE Symposium on*, pages 3–10. IEEE, 2013.

[Lim and Dey, 2010] Brian Y Lim and Anind K Dey. Toolkit to support intelligibility in context-aware applications. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 13–22. ACM, 2010.

[Lim *et al.*, 2009] Brian Y Lim, Anind K Dey, and Daniel Avrahami. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2119–2128. ACM, 2009.

[Luotsinen *et al.*, 2007] Linus J Luotsinen, Hans Fernlund, and Ladislau Bölöni. Automatic annotation of team actions in observations of embodied agents. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 9. ACM, 2007.

[MacFadden *et al.*, 2003] Alastair MacFadden, Lorin Elias, and Deborah Saucier. Males and females scan maps similarly, but give directions differently. *Brain and Cognition*, 53(2):297–300, 2003.

[Perera and Veloso, 2015] Vittorio Perera and Manuela Veloso. Handling Complex Commands for Service Robot Task Requests. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2015.

[Perera *et al.*, 2015] Vittorio Perera, Robin Soetens, Thomas Kollar, Mehdi Samadi, Yichao Sun, Daniele Nardi, René van de Molengraft, and Manuela Veloso. Learning task knowledge from dialog and web access. *Robotics*, 4(2):223–252, 2015.

[Rosenthal *et al.*, 2010] Stephanie Rosenthal, Joydeep Biswas, and Manuela Veloso. An Effective Personal Mobile Robot Agent Through Symbiotic Human-Robot Interaction. In *Proceedings of AAMAS'10, the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Toronto, Canada, May 2010.

[Thomason *et al.*, 2015] Jesse Thomason, Shiqi Zhang, Raymond Mooney, and Peter Stone. Learning to interpret natural language commands through human-robot dialog. In *Proceedings of the 2015 International Joint Conference on Artificial Intelligence (IJCAI)*, July 2015.

[Veloso *et al.*, 2008] Manuela Veloso, Nicholas Armstrong-Crews, Sonia Chernova, Elisabeth Crawford, Colin McMillen, Maayan Roth, Douglas Vail, and Stefan Zickler. A team of humanoid game commentators. *International Journal of Humanoid Robotics*, 5(03):457–480, 2008.

[Veloso *et al.*, 2015] Manuela Veloso, Joydeep Biswas, Brian Coltin, and Stephanie Rosenthal. CoBots: Robust Symbiotic Autonomous Mobile Service Robots. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2015.

[Voelz *et al.*, 1999] Dirk Voelz, Elisabeth André, Gerd Herzog, and Thomas Rist. Rocco: A robocup soccer commentator system. In *RoboCup-98*, pages 50–60. Springer, 1999.

[Yengui and Mahmoud, 2009] Ameni Yengui and NEJI Mahmoud. Semantic annotation formalism of video-conferencing documents. *Cognition and Exploratory Learning in Digital Age*, 2009.