Navigating Like a Human: Generating Landmark Designations for an Autonomous System Using a Novelty Algorithm

by Troy Dale Kelley and Sean M McGhee

Approved for public release; distribution is unlimited.
NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.
Navigating Like a Human: Generating Landmark Designations for an Autonomous System Using a Novelty Algorithm

Troy Dale Kelley
*Human Research and Engineering Directorate, ARL*

Sean M McGhee
*SOSi LLC, Reston, VA*
Navigating like a Human: Generating Landmark Designations for an Autonomous System using a Novelty Algorithm

To improve robotic navigation, and drawing on inspiration from human navigation, a landmark designation algorithm was developed and tested using a simulated environment and in the real world. The algorithm used laser data or visual data to determine whether a new environment was considered novel (i.e., different from the previous data), then used this designation for a landmark. The landmark algorithm was tested to determine 1) if it consistently designated the same locations as landmarks and 2) if landmark designation was sparse enough for practical use. Results showed that the algorithm generated relatively consistent landmarks that also corresponded to human landmark designations, but it could be improved by adding top-down contextual information to be more useful.
# Contents

List of Figures

iv

List of Tables

iv

1. Introduction

1

2. General Algorithm

5

2.1 Novelty Algorithm

5

2.2 Example Histogram

6

3. Testing Procedure

7

4. Results

8

5. Conclusions

10

6. References

11

List of Symbols, Abbreviations, and Acronyms

14

Distribution List

15
List of Figures

Fig. 1  Example of a novel interval .................................................................7
Fig. 2  Example of a bored interval .................................................................7
Fig. 3  Landmark designations autonomously identified on a simple map. Locations 1 and 2 were not marked, while Location 3 was marked correctly. ..............................................................................................................8
Fig. 4  Landmark designations autonomously identified for six locations on a complex map ........................................................................................................................................9

List of Tables

Table 1  Percentage of times the robot classified each of the six landmark locations across 14 map runs .........................................................................................................................10
1. Introduction

The US Army Research Laboratory’s (ARL’s) Human Research and Engineering Directorate began developing the Symbolic and Subsymbolic Robotics Intelligence Control System (SS-RICS) in 2006 (Kelley 2006). The SS-RICS is a production, goal-oriented, robotics control system based on the Adaptive Character of Thought–Rational (ACT-R) cognitive architecture (Anderson and Lebiere 2014). Additionally, development of the SS-RICS has leveraged heavily from biological and human capabilities for navigation, action selection, and memory decay. As part of the SS-RICS development, we were interested in looking at human-like forms of knowledge representation for navigation. Research has shown that humans “represent space in several independent, though interrelated, ways” (O'keefe and Nadel 1978) and that this representation is more robust and adaptive than the typical Artificial Intelligence (AI) algorithms used for spatial representation.

Navigation is critically important for Soldiers as well as for autonomous systems. For a robot to be useful on the battlefield, an autonomous system must have up-to-date knowledge about its current position in relationship to Soldiers, other robots, buildings, and so forth. While navigation seems to be an easy task for people to accomplish, it has remained an elusive one for autonomous systems researchers.

Navigation in AI is called Simultaneous Localization and Mapping (SLAM). A robot able to completely navigate a new environment has been called the “holy grail” of autonomous systems due to the difficulty and multifaceted nature of the task (Durrant-Whyte and Bailey 2006). The development of various SLAM algorithms has been an active area of research within the autonomous systems community for the past few decades (Imani et al. 2018) but is still considered an unresolved problem. As noted by Agunbiade and Zuva (2018), “there are still setbacks limiting the full acceptance of this technology even though the research had been conducted over the last 30 years.”

Several problem areas with the general SLAM technique have limited its usefulness for autonomous systems navigation. These problem areas have been identified in the literature as 1) loop closure, 2) correspondence, 3) dynamic mapping, and 4) kidnapping.

The first problem with the SLAM technique is known as “loop closure” (Glover et al. 2012). When loop closure fails to occur, the robot cannot determine if it has been to a previously visited location and cannot correctly complete the closure for a newly developed map (Ho and Newman 2007). Loop closure is also very scale-dependent. The larger the space covered during development of the map before loop closure occurs, the more likely it is that the map will not consolidate
correctly during loop closure, which can cause an error in map creation. Recent work to overcome problems associated with large scale map creation and loop closure errors has been done (Labbe and Michaud 2014), but the issue is not completely solved.

Another difficulty with the SLAM methodology is the so-called “correspondence problem” (Batlle et al. 1998). This problem, which occurs when a robot cannot determine if a new scene is completely novel or is the same as a previous scene, usually leads to localization errors within a SLAM map. The correspondence problem is related to a general issue with robotic perception, where the perceptual processes on a typical robotics system are not capable of accounting for dynamic changes in visual scenes (Kelley et al. 2011). Recent work to overcome both the correspondence and loop closure problems uses a visual vocabulary, or a “bag of words” algorithm, which represents a space as a histogram of words (Newman et al. 2006).

The correspondence problem is also related to the kidnapping problem (Labbe and Michaud 2014), which occurs when a robot is picked up or somehow moved in a way that does not allow it to simultaneously reference a SLAM map during the relocation process. As with the correspondence problem, when the robot is placed in the new location, it cannot determine that location in relationship to others on the SLAM map. Sometimes the SLAM algorithms will simply place the robot in the wrong location, which causes it to run into obstacles, since they are not shown on the SLAM map.

Finally, SLAM relies on certain basic (and sometimes erroneous) assumptions, one of the most prevalent being that nothing has changed in the world. In other words, the current map must be a static representation of the world in order for SLAM to localize correctly (Scona et al. 2018). If changes in the real world (i.e., a door is closed or obstacles are placed in a hallway) are not reflected on the static map, the robot can become disoriented or lost. Obviously, this has enormous implications for any use of SLAM on the battlefield, where dynamic changes are commonplace, and are indeed a basic assumption for a Warfighter. Any changes in the real world that are not reflected in a SLAM map will necessitate the creation of an entirely new SLAM map.

It would appear that many of these problems are caused by perceptual limitations in the sensory capabilities of autonomous systems. However, it could be that they are caused by limitations in the knowledge representation of a mapped environment and addressed by answering the following questions:

1) Can an efficient map of the environment be represented as a knowledge structure to overcome typical mapping problems?
2) Can efficient representations of space overcome the perceptual limitations inherent in noisy sensory data typically associated with robotic sensors that lead to SLAM mapping errors?

Humans generate a cognitive map of the world, a function commonly attributed to the hippocampus (O'keefe and Nadel 1978). The exact nature of this representation is still being debated by researchers (Warren et al. 2017), but is generally thought to be a topological map (Kuipers and Byun 1991; Thrun 1998; Yeap and Jefferies 1999; Kuipers et al. 2000; Kuipers et al. 2003), with large individual differences in accuracies (Weisberg and Newcombe 2018).

A topological map represents landmarks as nodes in a graph structure, and a “labeled graph” denotes distances and directions as arcs between the nodes (Warren et al. 2017). A topological map is a relational map, where space is represented as the relationship between material objects, and is in contrast to an absolute theory, which is a stationary framework for spatial representation (O'keefe and Nadel 1978) generally used in AI (i.e., SLAM). Topological maps enable more dynamic and robust representations than typical metric-based SLAM maps.

The advantage of a topological map is that it can handle dynamic changes between landmarks, as long as the landmarks remain in existence and are recognizable. For instance, lighting effects might cause the scenery between two locations to change, but the general locations and distances of the landmarks stay the same. Furthermore, a topological map can be more resistant to noise in sensor readings if the topological distances are assumed to be estimates of actual distances and the nodes are relative to each other in terms of direction. In other words, errors in distance measurements caused by wheel slippage for an autonomous system will have a negligible effect on a topological map if the distances between the landmarks are assumed to be estimates.

The term “landmark” is a new one for robotics, and consensus as to what constitutes a landmark has not been reached in the robotics literature, since any part of a SLAM map can be used as a landmark. In general, something that is consistently identifiable (e.g., the corner of a computer monitor) is considered a landmark according to robotics literature. In contrast, Gillner et al. (2008) note that human-based landmark designation tends to include three important criteria: 1) Salience—objects that are easily recognizable, 2) Relevance—objects that affect navigation decisions, and 3) Permanence—objects whose locations are not likely to change. The computer monitor corner mentioned earlier might not be salient if the lighting in the office has changed, and it is not relevant to general navigation.
within a building. Finally, someone might move the monitor, so it would not be considered permanent.

Semantically, humans think of a front door as a landmark for the interior of a building because it is a main entry and exit point and is “significant.” Whether or not a robot understands that a front door is the main entry and exit point of a building requires additional semantic understanding. If the robot understood this information, it might designate a front door as a landmark, irrespective of the semantic significance of a front door. Our landmark designation algorithm does not take a building’s semantics, design, or purpose into account. This is an obvious next step for our algorithm development, but was not pursued in this work.

Other researchers have used various algorithms to autonomously determine landmarks. For example, Jefferies et al. (2007) used Absolute Space Representations (ASRs) to determine landmark designations. To designate a landmark using an ASR, the algorithm “emphasizes the importance of detecting the exits in view from the surfaces perceived and from these exits a boundary for the local space is computed”. However, when doing this, the algorithm is metric space-oriented and is still susceptible to the metric errors associated with overlapping boundaries. Additionally, complex environments can “give rise to several ASRs” (Jefferies et al. 2007), which would lead to contradictory or overlapping landmarks.

Novelty algorithms have been used in the past for a variety of autonomous behaviors, including fault detection by robots used for inspections (Marsland et al. 2005). In general, novelty seems to be an important component of cognition (Borisyuk and Kazanovich 2004). It can attract attention (Itti and Baldi 2006), so can be used to trigger different behaviors. Novelty can be seen as a concept similar to “context updating” or the P300, which is ubiquitous in the electroencephalogram (EEG) literature (Donchin and Coles 1988). We use novelty in this work to trigger the designation of a landmark.

Our novelty algorithm for landmark designation defines a landmark as a significant shift from previous to new sensory information. (As noted earlier, semantics are not considered when designating a landmark.) Our novelty algorithm only interprets changes in space and does not rely on metric data to the point that it produces errors associated with overlapping boundary regions (Jefferies et al. 2007).
2. General Algorithm

The novelty algorithm was inspired by the orienting reflex, which was first discovered by Ivan Sechenov in 1863. To determine novelty, we assess how much variance there is in a sequence of consecutive sensory scan readings (equivalent to a time interval), specified by an interval parameter. If the variance does not meet certain criteria, the environment is not considered novel. In the case described in this report, we have applied this algorithm to laser scan readings (LIDAR, 180 points) but in our laboratory, we have applied it to images and auditory data as well.

Each scan is correlated with the others in the sequence, and those values are collected in a histogram matrix. If the correlation value of a particular comparison is greater than a set gradient value, then the two scenes are considered as the same and the count of correlation values used to calculate novelty is incremented. The gradient value determines the amount of change required to trigger a potential detection of a novel environment. A low value means that a great deal of change is necessary; a high value means that a relatively small amount of change is necessary. If the quotient exceeds the value in the parameter threshold, then the robot can be considered as not novel for that interval. Essentially, the more observations that are highly correlated within an interval, the more likely it is that the robot is seeing the same thing repeatedly and the more likely it is that the robot can be classified as not novel.

2.1 Novelty Algorithm

This section describes the novelty algorithm we used for landmark designation in detail. The algorithm uses observations or slices of data within a specified window, which are correlated and put into a histogram to determine a final threshold value.

Let $\alpha =$ vector of observations, 
$\beta =$ Interval of observations in the vector $\alpha$, 
$\mu =$ matrix of observations correlations (two dimensions), 
$T =$ Gradient value for a correlation value to be considered “novel”, 
$B =$ Threshold percentage of $T$ for the robot to be considered “novel” for the given set of observations, and 
$\gamma =$ number of correlations that exceed the novelty gradient $T$

\[
\gamma \leftarrow 0 \\
\text{for } i = 0 \rightarrow \beta - 1 \\
\quad \text{for } j = i+1 \rightarrow \beta - 1 \\
\quad \quad \mu_{ij} \leftarrow \text{correlation}(\alpha_i, \alpha_j) \\
\quad \quad \text{if } \mu_{ij} > T \text{ then } \gamma \leftarrow \gamma + 1 \\
\text{end} \\
\text{end}
\]
or, where \( f(x) = \text{correlation}(\alpha_i, \alpha_j) \):

\[
\gamma = \sum_{\substack{i,j=0 \ \text{to} \ \beta-1 \ \text{in} \ j \text{ greater than} \ i}} f(x) > T
\]  

(1)

Let \( \tau = \text{Threshold} \) percentage of correlations that exceed the novelty gradient \( T \):

\[
\tau = \frac{\gamma}{(\beta^2 - \beta)/2}.
\]  

(2)

if \( \tau > B \) then

RobotStatus ← “Not NOVEL”

else

RobotStatus ← “NOVEL”

end

2.2 Example Histogram

The three parameters explained in Section 2.1 were given the following values:

- **Gradient** = 99 (the correlation between observations must exceed .99 to be considered “not novel”)

- **Interval** = 50 (each assessment period will consist of 50 consecutive readings)

- **Threshold** = 75 (at least 75% of the correlations must exceed the “not novel” gradient for the interval to be considered a period of “not novel”)

From \((\text{Interval}^2 - \text{Interval}) / 2\), where \(\text{Interval} = 50\), we generated 1225 correlations that were used in the calculations. The histogram algorithm generates bins of 10 equally spaced intervals. Figure 1 shows that, of 1225 total correlations, approximately 420 can be considered “boring” (at or higher than the gradient value of 99 [last bar on the chart]). When plugged into Eq. 2, the boredom quotient is 0.34, which is lower than the Threshold value of 0.75; thus, the robot during this interval can be considered novel. In general, the more Gaussian the distribution of correlation values, the more novel the data.
Figure 2 shows that, of 1225 total correlations, there were approximately 1176 correlations that can be considered “boring” (at or higher than the Gradient value of 99). When plugged into Eq. 2, the boredom quotient is 0.96, which is higher than the Threshold value of 0.75; thus, the robot during this interval can be considered “bored”. In general, the more skewed the distribution, or the more correlations approaching a value of one, the less novel the data.

3. Testing Procedure

We tested our algorithm by using a Frontier Points algorithm (Yamauchi 1998) to explore a map until no Frontier Points were found. The novelty algorithm was continuously looking for novel locations while the Frontier Points algorithm was exploring the space. We ran the robot in a simple map in simulation (see Fig. 1), then 14 times in a more complex map (see Fig. 2), which was used for additional algorithm testing and improvements. Human researchers marked areas (generally
doorways) in the complex map as landmarks, and the robot was tested to find out if it would identify those same locations as landmarks.

4. Results

Figure 3 shows the landmark identification algorithm applied to a simple map in simulation for a single run. The small black dots are areas identified by the algorithm as novel. The robot is shown as a purple dot with a line extending from it. Its placement on the map shows where it finished its run using the Frontier Points exploration algorithm. As shown in the figure, the algorithm identified important transition areas on the map, including the entrances and exits from hallways and the entrances to rooms, but missed some important locations, such as Location 1 and Location 2. Additionally, the algorithm identified multiple instances in the same location (Location 3).

![Figure 3](image.png)

Fig. 3  Landmark designations autonomously identified on a simple map. Locations 1 and 2 were not marked, while Location 3 was marked correctly.

The algorithm uses three distance readings from the laser, which we used to modify it so that it would not mark the same location multiple times. We designated 0°–60° as left, 61°–120° as front, and 121°–180° as right. The largest reading was used as a classification of the area. For example, if the robot was in a hallway, the front would give the largest reading, and that area was classified as “front”. When a novel location was identified, the algorithm checked these previous sensor classifications to ensure the robot was in a new area (i.e., not still in the same area). If the sensor classification was different from the previous classification, the area was marked as novel. This modification to the algorithm reduced the number of
landmarks identified in a single location, as well as the overall number of landmarks identified.

Figure 4 shows landmarks identified by the novelty algorithm in a complex real-world map for a single run. The algorithm was modified to keep it from marking the same general location multiple times, which worked in some cases (Location 5) but not in others (Location 1). We were also interested in knowing whether or not the algorithm would identify as landmarks areas previously identified by humans as such. These were the major entrances and exits from rooms and hallways identified as Locations 1–6 in Fig. 4. As shown, the algorithm missed certain important areas (Location 6) but managed to identify most of the other areas correctly (Locations 1–5). Location 3 was identified correctly, but the modified algorithm kept the robot from identifying it repeatedly through the doorway. The algorithm began the identification of Location 3 correctly, at the doorway entrance, but did not identify the middle of the doorway as an important location. Changing the classification from “front” to “doorway” is an ideal way to improve the algorithm.

Additionally, we ran the robot through multiple (14) runs of the map displayed in Fig. 4 to see if it consistently identified the six locations we selected. These locations, all doorways, would be considered as important by humans, since they represent transitions from rooms and hallways. The results are shown in Table 1.
A Pioneer DX robot was started in the same location for each map run. The robot’s speed was typical for a Pioneer DX—about 2 mph. All locations were marked online during the run, and no post processing was required to mark the locations.

As shown in Table 1, the robot did not identify any of the six locations 100% of the time, but all locations were identified more than 50% of the time.

### Table 1  Percentage of times the robot classified each of the six landmark locations across 14 map runs

<table>
<thead>
<tr>
<th>Locations</th>
<th>Percentage Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location 1</td>
<td>57%</td>
</tr>
<tr>
<td>Location 2</td>
<td>78%</td>
</tr>
<tr>
<td>Location 3</td>
<td>57%</td>
</tr>
<tr>
<td>Location 4</td>
<td>71%</td>
</tr>
<tr>
<td>Location 5</td>
<td>57%</td>
</tr>
<tr>
<td>Location 6</td>
<td>71%</td>
</tr>
</tbody>
</table>

5. Conclusions

We tested the consistency of a novelty algorithm, as an initial first pass, for landmark designation and for eventual use in a topological map. After initial testing, the algorithm was modified so that it did not mark the same location repeatedly. This modification worked well (see Fig. 4).

The algorithm marked areas for which drastic changes in space had occurred as landmarks. This also tended to correlate well with the beginnings of hallways and the exits of rooms, or regions that human researchers had also designated as landmarks. We found mixed results in consistency across runs of the same map. For instance, when we tried to match the robot landmarks with human-based landmark designation across 14 runs of the same real-world map, the robot marked one location 78% of the time, but there were locations that it only marked 57% of the time.

Semantically, it is necessary for the robot to understand that doorways are important and to be able to identify them as such. This would further improve the matching between the human and robot landmark designations. The novelty algorithm could be used as a cue to look for a doorway or other important or salient structures in the environment. Further work needs to be done to develop a topological map using the novelty-generated landmark designations.
6. References


### List of Symbols, Abbreviations, and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT-R</td>
<td>Adaptive Character of Thought–Rational</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>ARL</td>
<td>US Army Research Laboratory</td>
</tr>
<tr>
<td>ASR</td>
<td>Absolute Space Representations</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalogram</td>
</tr>
<tr>
<td>SLAM</td>
<td>Simultaneous Localization and Mapping</td>
</tr>
<tr>
<td>SS-RICS</td>
<td>Symbolic and Subsymbolic Robotics Intelligence Control System</td>
</tr>
</tbody>
</table>