A Logical Account of the Common Sense Informatic Situation for a Mobile Robot

Murray Shanahan

Imperial College
Department of Electrical and Electronic Engineering,
Exhibition Road,
London SW7 2BT,
England.

Tel: +44 (0)171 594 6307
Email: m.shanahan@ic.ac.uk

January 1996
Last modified December 1998

Abstract

Any model of the world a robot constructs on the basis of its sensor data is necessarily both incomplete, due to the robot's limited window on the world, and uncertain, due to sensor and motor noise. This paper proposes a logic-based framework in which such models are constructed through an abductive process whereby sensor data is explained by hypothesising the existence, locations, and shapes of objects. Symbols appearing in the resulting explanations acquire meaning through the theory, and yet are grounded by the robot's interaction with the world. The proposed framework draws on existing logic-based formalisms for representing action, continuous change, space, and shape, but a novel solution to the frame problem is employed. Noise is treated as a kind of non-determinism, and is dealt with by a consistency-based form of abduction.[1]

Keywords: Cognitive Robotics, Common Sense Reasoning.

---

[1] Much of the work in this article was published in two conference papers in 1996, one in the ECAI conference [Shanahan, 1996a], where it won the Best Paper Award, and one in the AAAI conference [Shanahan, 1996b]. Since 1996, the author has carried out further work on the topic of the article, to which the interested reader may wish to refer, namely [Shanahan, 1997b] and [Shanahan, 1998].

Introduction

Since the end of the Shakey project [Nilsson, 1984], an ever widening gap has opened up between theoretical work in Knowledge Representation and the practice of robotics. At the end of the Eighties, this trend culminated in the work of Brooks, who rejected the assumptions underlying the Shakey project altogether [Brooks, 1986], [Brooks, 1991].

> The key problem I see with [all work in the style of Shakey] is that it relied on the assumption that a complete world model could be built internally and then manipulated.
>
> [Brooks, 1991a, page 577]

As Brooks points out, a complete model of the world is hard for a robot to construct because,

> The data delivered by sensors are not direct descriptions of the world as objects and their relationships [and] commands to actuators have very uncertain effects.
>
> [Brooks, 1991b, page 5]

Without ignoring the lessons of the early Seventies, the nascent area of Cognitive Robotics [Lespérance, *et al.*, 1994] seeks to reinstate the ideals of the Shakey project, namely the construction of robots whose architecture is based on the idea of representing the world by sentences of formal logic and reasoning about it by manipulating those sentences. The chief benefits of this approach are,

- that it facilitates the endowment of a robot with the capacity to perform high-level reasoning tasks, such as planning, and
- that it makes it possible to formally account for the success (or otherwise) of a robot by appealing to the notions of correct reasoning and correct representation.

Where Brooks and his followers see residual adherence to the ideals of the Shakey project as a barrier to progress in AI, Cognitive Robotics laments the failure to pursue those ideals with sufficient zeal. Contrary to ill-informed opinion, logic is eminently suited to representing incomplete models of the world, which are the only models that can be built on the basis of the incomplete and noisy data delivered by a robot's sensors.

Unlike the Shakey project, contemporary Cognitive Robotics can draw on twenty five years of progress in the formalisation of common sense reasoning, specifically on representing and reasoning about action, change, space, and shape, and especially on reasoning with incomplete information. In addition, it benefits from advances in a number of enabling technologies, including logic programming techniques, algorithms from mainstream robotics, and low-cost mobile robotic platforms with substantial on-board computing power.

This paper concerns the representation of knowledge about the objects in a robot's environment, and how such knowledge is acquired. The main feature of this knowledge is its incompleteness and uncertainty, placing the robot in what McCarthy calls the *common sense informatic situation* [1989]. The treatment given in the paper is rigorously logical, but has been carried through to implementation on a real robot.

The paper is organised as follows. After outlining the basic idea of abductive sensor data assimilation, a general purpose formalism for representing action, continuous change, space and shape is introduced. This formalism is then used to represent the relationship between the actions of a particular (real) robot and the sensor data it receives from the world. The abductive task of assimilating sensor data is then straightforwardly characterised. The next topic to be considered is that of noise, which is taken to mean uncertainty in both sensor input and motor output. Two theorems are then developed which facilitate the design of map-building algorithms that can be executed on the robot. Finally, the results of preliminary experimentation with the robot are reported.

1 Assimilating Sensor Data

The key idea of this paper is to consider the process of assimilating a stream of sensor data as abduction. Given such a stream, the abductive task is to hypothesise the existence, shapes, and locations of objects which, given the output the robot has supplied to its motors, would explain that sensor data [Charniak & McDermott, 1985, page 455]. This is, in essence, the map building task for a mobile robot.

More precisely, if a stream of sensor data is represented as the conjunction Ψ of a set of observation sentences, the task is to find an explanation of Ψ in the form of a logical description (a map) Δ_M of the initial locations and shapes of a number of objects, such that,

$$\Sigma_B \wedge \Sigma_E \wedge \Delta_N \wedge \Delta_M \models \Psi$$

where,

- Σ_B is a background theory, comprising axioms for change (including continuous change), action, space, and shape,
- Σ_E is a theory relating the shapes and movements of objects (including the robot itself) to the robot's sensor data, and
- Δ_N is a logical description of the movements of objects, including the robot itself.

The exact form of these components is described in the next three sections, which present formalisms for representing and reasoning about action, change, space, and shape. In practice, as we'll see, these components will have to be split into parts for technical reasons.

Three major issues arise with this logical specification of the map building task: noisy data, incomplete information, and implementation.

- Σ_E does not have to assume a perfect correspondence between objects in the world and sensor data received from them, or a perfect correspondence between motor outputs and actual movements in the world. In practice, a noisy interface between world and robot must be assumed. Using the expressive power of first-order logic, the uncertainty resulting from such noise can be captured.
- Data in the common sense informatic situation is incomplete as well as noisy. In abductive terms, there will typically be many Δ_M's that could explain any given Ψ. For example, the robot may only receive sensor data from a small fraction of the total surface of an object, and be unable to tell whether the object is large or small. Again, using the expressive power of first-order logic, this incompleteness can be captured.
- This logical specification of the map building task must be rendered into an efficient implementation which can be executed by the on-board microprocessor of a mobile robot.

The provision of a logic-based theoretical account brings issues like noise and incompleteness into sharp focus, and permits their study within the same framework used to address wider epistemological questions in knowledge representation. It also enables the formal evaluation of algorithms for low-level motor-perception tasks by supplying a formalism in which these tasks can be precisely specified. These topics are discussed in detail in later sections.

2 Representing Action

The formalism used in this paper to represent action and change, including continuous change, is adapted from the circumscriptive event calculus presented in [Shanahan, 1995b], which in turn is based loosely on the formalism of Kowalski and Sergot [1986]. However, it employs a novel solution to the frame problem, inspired by the work of Kartha and Lifschitz [1995]. The result is a considerable simplification of the formalism in [Shanahan, 1995b].

Throughout the paper, the language of many-sorted first-order predicate calculus with equality will be used, augmented with circumscription [McCarthy, 1986], [Lifschitz, 1994]. Variables in formulae begin with lower-case letters and are universally quantified with maximum scope unless indicated otherwise.

In the event calculus, we have sorts for *fluents*, *actions* (or events), and *time points*. It's assumed that time points are interpreted by the reals, and that the usual comparative predicates, arithmetic functions, and trigonometric functions are suitably defined. The formula HoldsAt(f,t) says that fluent f is true at time point t. The formulae Initiates(a,f,t) and Terminates(a,f,t) say respectively that action a makes fluent f true from time point t, and that a makes f false from t. The effects of actions are described by a collection of formulae involving Initiates and Terminates.

For example, if the term Rotate(r) denotes a robot's action of rotating r degrees about some axis passing through its body, and the term Facing(r) is a fluent representing that the robot is facing in a direction r degrees from North, then we might write the following Initiates and Terminates formulae.[2]

$$\text{Initiates}(\text{Rotate}(r1),\text{Facing}(r2),t) \leftarrow \text{HoldsAt}(\text{Facing}(r3),t) \wedge r2 = r3 + r1 \tag{2.1}$$

$$\text{Terminates}(\text{Rotate}(r1),\text{Facing}(r2),t) \leftarrow \text{HoldsAt}(\text{Facing}(r2),t) \wedge r1 \neq 0 \tag{2.2}$$

Once a fluent has been initiated or terminated by an action or event, it is subject to the common sense law of inertia, which is captured by the event calculus axioms to be presented shortly. This means that it retains its value (true or false) until another action or event occurs which affects that fluent.

A narrative of actions and events is described via the predicates Happens, Initially_P, and Initially_N. The formula Happens(a,t) says that an action or event of type a occurred at time point t. Events are instantaneous. The formula $\text{Initially}_P(f)$ says that the fluent f is true from time point 0, and the formula $\text{Initially}_N(f)$ says that the fluent f is false from time point 0. Here's an example narrative.

$$\text{Initially}_P(\text{Facing}(0)) \tag{2.3}$$

$$\text{Happens}(\text{Rotate}(90),10) \tag{2.4}$$

$$\text{Happens}(\text{Rotate}(-180),20) \tag{2.5}$$

A theory will also include a pair of uniqueness-of-names axioms, one for actions and one fluents.[3]

$$\text{UNA}[\text{Facing}] \tag{2.6}$$

$$\text{UNA}[\text{Rotate}] \tag{2.7}$$

The relationship between HoldsAt, Happens, Initiates, and Terminates is constrained by the following axioms. Note that a fluent does not hold at the time of an action or event that initiates it, but does hold at the time of an action or event that terminates it.

$$\text{HoldsAt}(f,t) \leftarrow \text{Initially}_P(f) \wedge \neg\, \text{Clipped}(0,f,t) \tag{EC1}$$

$$\begin{aligned}&\text{HoldsAt}(f,t2) \leftarrow \\ &\quad \text{Happens}(a,t1) \wedge \text{Initiates}(a,f,t1) \wedge t1 < t2 \wedge \neg\, \text{Clipped}(t1,f,t2)\end{aligned} \tag{EC2}$$

$$\neg\, \text{HoldsAt}(f,t) \leftarrow \text{Initially}_N(f) \wedge \neg\, \text{Declipped}(0,f,t) \tag{EC3}$$

$$\begin{aligned}&\neg\, \text{HoldsAt}(f,t2) \leftarrow \\ &\quad \text{Happens}(a,t1) \wedge \text{Terminates}(a,f,t1) \wedge t1 < t2 \wedge \neg\, \text{Declipped}(t1,f,t2)\end{aligned} \tag{EC4}$$

---

[2] Rotation is treated as instantaneous here, and throughout the sequel. So long as the robot always finishes turning before it moves forwards, this idealisation is acceptable.
[3] The UNA notation is taken from [Baker, 1989].

$\text{Clipped}(t1,f,t2) \leftrightarrow$ (EC5)
$$\exists\, a,t\ [\text{Happens}(a,t) \wedge [\text{Terminates}(a,f,t) \vee \text{Releases}(a,f,t)] \wedge t1 < t \wedge t < t2]$$

$\text{Declipped}(t1,f,t2) \leftrightarrow$ (EC6)
$$\exists\, a,t\ [\text{Happens}(a,t) \wedge [\text{Initiates}(a,f,t) \vee \text{Releases}(a,f,t)] \wedge t1 < t \wedge t < t2]$$

These axioms introduce a new predicate Releases [Kartha & Lifschitz, 1994]. The formula Releases(a,f,t) says that action a exempts fluent f from the common sense law of inertia.[4] This non-inertial status is revoked as soon as the fluent is initiated or terminated once more. The use of this predicate will be illustrated shortly in the context of continuous change.[5]

Let the conjunction of (EC1) to (EC6) be denoted by EC. The circumscription policy to overcome the frame problem is the following. Given a conjunction N of Happens, Initially_P, and Initially_N formulae, a conjunction E of Initiates, Terminates and Releases formulae, and a conjunction U of uniqueness-of-names axioms, we are interested in,

$$\text{CIRC}[N\ ;\ \text{Happens}] \wedge \text{CIRC}[E\ ;\ \text{Initiates, Terminates, Releases}] \wedge U \wedge EC$$

This formula embodies a form of the common sense law of inertia, and thereby solves the frame problem. To see why in detail, consult [Shanahan, 1997a]. The key to the solution is to put EC outside the scope of the circumscriptions, thus ensuring that the Hanks-McDermott problem is avoided [Hanks & McDermott, 1987].[6] In most cases, the two circumscriptions will yield predicate completions, making the overall formula manageable and intuitive.

For the example above, we have the following proposition. Let E be the conjunction of (2.1) with (2.2), let N be the conjunction of (2.3) to (2.5), and let U be the conjunction of (2.6) with (2.7).

Proposition 2.8.

$$\text{CIRC}[N\ ;\ \text{Happens}] \wedge \text{CIRC}[E\ ;\ \text{Initiates, Terminates, Releases}] \wedge U \wedge EC \vDash$$
$$\text{HoldsAt}(\text{Facing}(r),t) \leftarrow [0 \leq t \leq 10 \wedge r = 0] \vee [10 < t \leq 20 \wedge r = 90] \vee [20 < t \wedge r = 270].$$

Proof. See Appendix A. □

3 Domain Constraints and Continuous Change

Two additional features of the calculus are important: the ability to represent domain constraints, and the ability to represent continuous change.

---

[4] The Releases predicate is related to Sandewall's idea of occlusion [Sandewall, 1994].
[5] Formula (2.2) above can be written using Releases instead of Terminates, so long as there is a domain constraint ensuring that the robot always has a unique orientation.
[6] This is related to Sandewall's idea of filtered preferential entailment [Sandewall, 1994].

Domain constraints are straightforwardly dealt with in the proposed formalism. They are simply formulated as HoldsAt formulae with a single universally quantified time variable, and conjoined outside the scope of the circumscriptions along with EC. For example, the following domain constraint expresses the fact that the robot can only face in one direction at a time.

$$\text{HoldsAt(Facing(r1),t)} \wedge \text{HoldsAt(Facing(r2),t)} \rightarrow \text{r1} = \text{r2}$$

In the event calculus, domain constraints are used to determine values for fluents that haven't been initiated or terminated by actions or events (non-inertial fluents) given the values of other fluents that have. (Domain constraints that attempt to constrain the relationship between inertial fluents can lead to inconsistency.)[7]

The issue of continuous change has been largely neglected in the design of formalisms for reasoning about action until recently [Sandewall, 1989], [Shanahan, 1990], [Miller, 1996], [Reiter, 1996]. In the present formalism, following [Shanahan, 1990], continuous change is represented through the introduction of a new predicate and the addition of an extra axiom. The formula Trajectory(f1,t,f2,d) represents that, if the fluent f1 is initiated at time t, then after a period of time d the fluent f2 holds. We have the following axiom.

$$\begin{aligned}&\text{HoldsAt(f2,t2)} \leftarrow \\ &\quad \text{Happens(a,t1)} \wedge \text{Initiates(a,f1,t1)} \wedge \text{t1} < \text{t2} \wedge \text{t2} = \text{t1} + \text{d} \wedge \\ &\quad\quad \text{Trajectory(f1,t1,f2,d)} \wedge \neg\, \text{Clipped(t1,f1,t2)}\end{aligned} \tag{EC7}$$

Let CEC denote EC $\wedge$ (EC7), and U denote the conjunction of a set of uniqueness-of-names axioms. If R is the conjunction of a set of domain constraints and T is the conjunction of set of formulae constraining Trajectory, then we are interested in,

$$\begin{aligned}&\text{CIRC[N ; Happens]} \wedge \\ &\quad \text{CIRC[E ; Initiates, Terminates, Releases]} \wedge \\ &\quad\quad \text{T} \wedge \text{R} \wedge \text{U} \wedge \text{CEC}.\end{aligned}$$

For example, suppose the robot's repertoire of actions is expanded to include the actions Go and Stop. The Go action initiates a period of continuous change in the robot's location. The Stop action terminates such a period. For the present, the robot's location will be represented by the fluent Location(Robot,p), where p is a pair of Cartesian co-ordinates the form $\langle x,y \rangle$.[8] A constant velocity V is assumed in the following collection of formulae, which are intended to capture this example.

Let E be the conjunction of the following formulae.

$$\text{Initiates(Go,Moving,t)} \tag{3.1}$$

$$\text{Releases(Go,Location(Robot,p),t)} \tag{3.2}$$

7 Note that Initiates(a,F1,t) $\rightarrow$ Initiates(a,F2,t) does not follow from HoldsAt(F1,t) $\rightarrow$ HoldsAt(F2,t).

8 Later, the Location fluent will be replaced by a fluent which captures spatial occupancy.

$$\text{Terminates(Stop,Moving,t)} \tag{3.3}$$

$$\text{Initiates(Stop,Location(Robot,p),t)} \leftarrow \text{HoldsAt(Location(Robot,p),t)} \tag{3.4}$$

Let T be the following formula.

$$\begin{aligned}&\text{Trajectory(Moving,t,Location(Robot,}\langle x2,y2\rangle\text{),d)} \leftarrow \\ &\quad \text{HoldsAt(Location(Robot,}\langle x1,y1\rangle\text{),t)} \wedge \text{HoldsAt(Facing(r),t1)} \wedge \\ &\qquad x2 = x1 + V.d.\mathrm{Sin}(r) \wedge y2 = y1 + V.d.\mathrm{Cos}(r)\end{aligned} \tag{3.5}$$

Let R be the following domain constraint.

$$\text{HoldsAt(Location(w,p1),t)} \wedge \text{HoldsAt(Location(w,p2),t)} \rightarrow p1 = p2 \tag{3.6}$$

Let U be the conjunction of the following uniqueness-of-names axioms.

$$\text{UNA[Location, Facing, Moving]} \tag{3.7}$$

$$\text{UNA[Go, Stop]} \tag{3.8}$$

Let N be the following narrative description.

$$\text{Initially}_\text{P}\text{(Location(Robot,}\langle 0,0\rangle\text{)} \tag{3.9}$$

$$\text{Initially}_\text{P}\text{(Facing(90))} \tag{3.10}$$

$$\text{Happens(Go,10)} \tag{3.11}$$

$$\text{Happens(Stop,20)} \tag{3.12}$$

Now, given that the circumscriptions of E and N yield the predicate completions of Happens, Initiates, Terminates, and Releases, it’s a straightforward exercise to show that the recommended circumscription yields what we would expect.

Proposition 3.13.

$$\begin{aligned}&\text{CIRC[N ; Happens]} \wedge \\ &\quad \text{CIRC[E ; Initiates, Terminates, Releases]} \wedge \text{T} \wedge \text{R} \wedge \text{U} \wedge \text{CEC} \vDash \\ &\qquad \text{HoldsAt(Location(Robot,}\langle x,y\rangle\text{),t)} \leftrightarrow \\ &\qquad\quad [0 \leq t \leq 10 \wedge x = 0 \wedge y = 0] \vee \\ &\qquad\qquad [10 < t \leq 20 \wedge x = V.(t - 10) \wedge y = 0] \vee \\ &\qquad\qquad\quad [20 < t \wedge x = V.10 \wedge y = 0].\end{aligned}$$

Proof. See Appendix A. □

Notice that we are at liberty to include formulae which describe triggered events in N. Here’s an example of such a formula, which describes conditions under which the robot will collide with a wall lying on an East-West line 100 units north of the origin.

$$\begin{aligned}&\text{Happens(Bump,t)} \leftarrow \\ &\quad \text{HoldsAt(Moving,t)} \wedge \text{HoldsAt(Facing(r),t)} \wedge \\ &\qquad -90 < r < 90 \wedge \text{HoldsAt(Location(Robot,}\langle x,90\rangle\text{),t)}\end{aligned}$$

4 Representing Space and Shape

The formalism used in this paper to represent space and shape is adapted from [Shanahan, 1995a].[9] Space is considered a real-valued co-ordinate system.[10] For present purposes we can take space to be the plane $\mathbb{R} \times \mathbb{R}$, reflecting the fact that the robot under consideration only moves in two dimensions. A *region* is a subset of $\mathbb{R} \times \mathbb{R}$. A *point* is a member of $\mathbb{R} \times \mathbb{R}$. I will consider only interpretations in which points are interpreted as pairs of reals, in which regions are interpreted as sets of points, and in which the $\in$ predicate has its usual meaning.

Objects occupy open, path-connected regions. For example, the following formula describes an open circle of radius z units centred on the origin.

$$\mathrm{p} \in \mathrm{Disc(z)} \leftrightarrow \mathrm{Distance(p,}\langle 0,0\rangle) < \mathrm{z} \qquad \text{(Sp1)}$$

Distance is a function yielding a positive real number, defined in the obvious way.

$$\mathrm{Distance}(\langle x1,y1\rangle,\langle x2,y2\rangle) = \sqrt{(x1-x2)^2 + (y1-y2)^2} \qquad \text{(Sp2)}$$

The function Bearing is also useful.

$$\begin{aligned} &\mathrm{Bearing}(\langle x1,y1\rangle,\langle x2,y2\rangle) = r \leftarrow \qquad \text{(Sp3)} \\ &\quad z = \mathrm{Distance}(\langle x1,y1\rangle,\langle x2,y2\rangle) \wedge z \neq 0 \wedge \\ &\qquad \mathrm{Sin}(r) = \frac{x2-x1}{z} \wedge \mathrm{Cos}(r) = \frac{y2-y1}{z} \end{aligned}$$

Using Distance and Bearing we can define a straight line as follows. The term Line(p1,p2) denotes the straight line segment whose end points are p1 and p2. The Line function is useful in defining shapes with straight line boundaries.

$$\begin{aligned} &p \in \mathrm{Line}(p1,p2) \leftrightarrow \qquad \text{(Sp4)} \\ &\quad \mathrm{Bearing}(p1,p) = \mathrm{Bearing}(p1,p2) \wedge \mathrm{Distance}(p1,p) \leq \mathrm{Distance}(p1,p2) \end{aligned}$$

Spatial occupancy is represented by the fluent Occupies. The term Occupies(w,g) denotes that object w occupies region g. No object can occupy two regions at the same time. This implies, for example, that if an object occupies a region g, it doesn't occupy any subset of g nor any superset of g. We have the following domain constraints.

$$\mathrm{HoldsAt}(\mathrm{Occupies}(w,g1),t) \wedge \mathrm{HoldsAt}(\mathrm{Occupies}(w,g2),t) \rightarrow g1 = g2 \qquad \text{(Sp5)}$$

$$\begin{aligned} &\mathrm{HoldsAt}(\mathrm{Occupies}(w1,g1),t) \wedge \qquad \text{(Sp6)} \\ &\quad \mathrm{HoldsAt}(\mathrm{Occupies}(w2,g2),t) \wedge w1 \neq w2 \rightarrow \\ &\qquad \neg \exists p\, [p \in g1 \wedge p \in g2] \end{aligned}$$

The first of these axioms captures the uniqueness of an object's region of occupancy, and the second insists that no two objects overlap.

---

[9] The Location fluent and Shape function from [Shanahan, 1995a] have been dispensed with.

[10] This decision does not in any way rule out the adoption of qualitative approaches to spatial reasoning, in the spirit of the Naive Physics Manifesto [Hayes, 1985]. The real-valued co-ordinate system can be thought of as underpinning these approaches.

The term Displace(g,$\langle x,y \rangle$) denotes the result of displacing the region g by x units east and y units north.

$$\langle x1,y1 \rangle \in \text{Displace}(g,\langle x2,y2 \rangle) \leftrightarrow \langle x1-x2,y1-y2 \rangle \in g \qquad \text{(Sp7)}$$

The Displace function is primarily used to describe motion: if an object moves, the region it occupies is displaced. But it can also be used to combine regions by taking their union (via disjunction). The following formula defines a shape a little like the field of view through a pair of binoculars, formed from two overlapping circles.

$$p \in \text{TwoDiscs}(x) \leftrightarrow p \in \text{Displace}(\text{Disc}(x),\langle -\frac{x}{2},0 \rangle) \vee p \in \text{Displace}(\text{Disc}(x),\langle \frac{x}{2},0 \rangle)$$

The final component of the framework is a means of default reasoning about spatial occupancy [Shanahan, 1995a]. Shortly, a theory of continuous motion will be described. This theory insists that, in order for an object to follow a trajectory in space, that trajectory must be clear. Accordingly, as well as capturing which regions of space are occupied, our theory of space and shape must capture which regions are unoccupied.

A suitable strategy is to make space empty by default. It's sufficient to apply this default just to the situation at time 0 — the common sense law of inertia will effectively carry it over to later times. The following axiom is required, which can be thought of as a *common sense law of spatial occupancy*.

$$\text{AbSpace}(w) \leftarrow \exists\, g\, [\neg\, \text{Initially}_N(\text{Occupies}(w,g))] \qquad \text{(Sp8)}$$

The predicate AbSpace needs to be minimised, with Initially_P and Initially_N allowed to vary. (Note that, from the event calculus axioms, for any w and g, $\text{Initially}_N(\text{Occupies}(w,g))$ implies $\neg\, \text{Initially}_P(\text{Occupies}(w,g))$.)

Where previously we were interested in CIRC[N ; Happens], it's now convenient to split this circumscription into two. Given,

- the conjunction O of Axioms (Sp1) to (Sp8),
- a conjunction M of Initially_P and Initially_N formulae which mention only the fluent Occupies, and
- a conjunction N of Happens formulae and Initially_P and Initially_N formulae which don't mention the fluent Occupies, and
- conjunctions E, T, R, U, and CEC as described in the last section,

we are now interested in,

$$\begin{aligned} &\text{CIRC}\,[O \wedge M\ ;\ \text{AbSpace}\ ;\ \text{Initially}_P, \text{Initially}_N] \wedge \text{CIRC}[N\ ;\ \text{Happens}] \wedge \\ &\quad \text{CIRC}[E\ ;\ \text{Initiates, Terminates, Releases}] \wedge \\ &\qquad T \wedge R \wedge U \wedge \text{CEC}. \end{aligned}$$

5 Sensors and Motors: The Theory Σ_E

We now have the logical apparatus required to construct a formal theory of the relationship between a robot's motor activity, the world, and the robot's sensor data. For now we will assume perfect motors and perfect sensors. The issue of noise is dealt with in Section 7.

The robot used as an example throughout the rest of the paper is one of the simplest and cheapest commercially available mobile robotic platforms at the time of writing, namely the Rug Warrior described by Jones and Flynn [1993] (Figure 5a). This is a small (approximately 20cm diameter), wheeled robot with a 68000 series microprocessor plus 32K RAM on board. It has a very simple collection of sensors. These include three bump switches arranged around its circumference, which will be our main concern here. In particular, we will confine our attention to the two forward bump switches, which, in combination, can deliver three possible values for the direction of a collision.

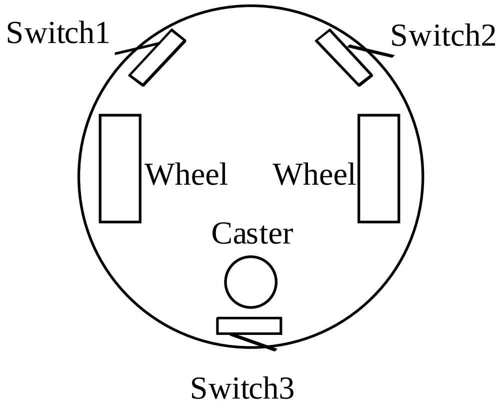

Figure 5a: The Rug Warrior Robot from Above

Needless to say, each different kind of sensor gives rise to its own particular set of problems when it comes to constructing Σ_E. The question of noise is largely irrelevant when it comes to bump sensors. With infra-red proximity detectors, noise plays a small part. With sonar, the significance of noise is much greater. The use of cameras gives rise to a whole set of issues which are beyond the scope of this paper.

The central idea of this paper is the assimilation of sensor data through abduction. This is in accordance with the principle, "prediction is deduction but explanation is abduction" [Shanahan, 1989]. To begin with, we'll be looking at the predictive capabilities of the framework described. The conjunction of our general theory of action, change, space, and shape with the theory Σ_E, along with a description of the initial locations and shapes of objects in the world and a description of the robot's actions, should yield a description of the robot's expected sensory input. If prediction works properly using deduction in this way, the reverse operation of explaining a given stream of sensor data by hypothesising the locations and shapes of objects in the world is already defined. It is simply abduction using the same logical framework.

In the caricature of the task of assimilating sensor data presented in Section 1, the relationship between motor activity and sensor data was described by Σ_E. In practice, this theory is split into parts and distributed across different circumscriptions (see Section 3).

First, we have a collection of formulae which are outside the scope of any circumscription. Let B be the conjunction of CEC with Axioms (B1) to (B6) below. The robot is assumed to travel at a velocity of one unit of distance per unit of time.

$$\text{UNA[Occupies, Facing, Moving, Blocked, Touching]} \tag{B1}$$

$$\text{UNA[Rotate, Go, Stop, Bump, Switch1, Switch2]} \tag{B2}$$

$$\begin{aligned}&\text{Trajectory(Moving,t,Occupies(Robot,g2),d)} \leftarrow \\ &\quad \text{HoldsAt(Occupies(Robot,g1),t)} \wedge \text{HoldsAt(Facing(r),t)} \wedge \\ &\quad\quad \text{g2} = \text{Displace(g1,}\langle\text{d.Sin(r),d.Cos(r)}\rangle\text{)}\end{aligned} \tag{B3}$$

$$\text{HoldsAt(Facing(r1),t)} \wedge \text{HoldsAt(Facing(r2),t)} \rightarrow \text{r1} = \text{r2} \tag{B4}$$

$$\begin{aligned}&\text{HoldsAt(Blocked(w1,w2,r),t)} \leftrightarrow \\ &\quad \exists\, \text{g1,g2 [HoldsAt(Occupies(w1,g1),t)} \wedge \text{HoldsAt(Occupies(w2,g2),t)} \wedge \\ &\quad\quad \text{w1} \neq \text{w2} \wedge \exists\, \text{z1 [z1} > 0 \wedge \forall\, \text{z2 [z2} \leq \text{z1} \rightarrow \\ &\quad\quad\quad \exists\, \text{p [p} \in \text{g2} \wedge \text{p} \in \text{Displace(g1,}\langle\text{z2.Sin(r),z2.Cos(r)}\rangle\text{)]]]]}\end{aligned} \tag{B5}$$

$$\begin{aligned}&\text{HoldsAt(Touching(w1,w2,p),t)} \leftrightarrow \\ &\quad \exists\, \text{g1,g2 [HoldsAt(Occupies(w1,g1),t)} \wedge \text{HoldsAt(Occupies(w2,g2),t)} \wedge \\ &\quad\quad \text{w1} \neq \text{w2} \wedge \exists\, \text{p1, p2 [p} \in \text{Line(p1,p2)} \wedge \text{p} \neq \text{p1} \wedge \text{p} \neq \text{p2} \wedge \\ &\quad\quad\quad \forall\, \text{p3 [[p3} \in \text{Line(p1,p)} \wedge \text{p3} \neq \text{p]} \rightarrow \text{p3} \in \text{g1]} \wedge \\ &\quad\quad\quad\quad \forall\, \text{p3 [[p3} \in \text{Line(p,p2)} \wedge \text{p3} \neq \text{p]} \rightarrow \text{p3} \in \text{g2]]]}\end{aligned} \tag{B6}$$

The fluent Blocked(w1,w2,r) holds if object w1 cannot move any distance at all in direction r without overlapping with another object, namely w2. The fluent Touching(w1,w2,p) holds if w1 and w2 are touching at point p. This is true if a straight line exists from p1 to p2 at a bearing r which includes a point p3 such that every point between p1 and p3 apart from p3 itself is in g1 and every point from p2 to p3 apart from p3 itself is in g2.

Next we have a collection of Initiates, Terminates, and Releases formulae. Let E be the conjunction of the following axioms (E1) to (E6). A Bump event occurs when the robot collides with something.

$$\text{Initiates(Rotate(r1),Facing(r1+r2),t)} \leftarrow \text{HoldsAt(Facing(r2),t)} \tag{E1}$$

$$\text{Releases(Rotate(r1),Facing(r2),t)} \leftarrow \text{HoldsAt(Facing(r2),t)} \wedge \text{r1} \neq 0 \tag{E2}$$

$$\text{Initiates(Go,Moving,t)} \tag{E3}$$

$$\text{Releases(Go,Occupies(Robot,g),t)} \tag{E4}$$

$$\text{Terminates(a,Moving,t)} \leftarrow \text{a} = \text{Stop} \vee \text{a} = \text{Bump} \vee \text{a} = \text{Rotate(r)} \tag{E5}$$

Initiates(a,Occupies(Robot,g),t) $\leftarrow$ (E6)
[a = Stop $\vee$ a = Bump] $\wedge$ HoldsAt(Occupies(Robot,g),t)

Now we have a collection of formulae concerning the narrative of actions and events we're interested in. This collection has two parts. Let N be N1 $\wedge$ N2. The first component part concerns triggered events. The events Switch1 and Switch2 occur when the robot's forward bump switches are tripped (see Figure 5a). Let N1 be the conjunction of Axioms (H1) to (H3) below.[11]

Happens(Bump,t) $\leftarrow$ (H1)
[HoldsAt(Moving,t) $\vee$ Happens(Go,t)] $\wedge$
HoldsAt(Facing(r),t) $\wedge$ HoldsAt(Blocked(Robot,w,r),t)

Happens(Switch1,t) $\leftarrow$ (H2)
Happens(Bump,t) $\wedge$ HoldsAt(Facing(r),t) $\wedge$
HoldsAt(Occupies(Robot,Displace(Disc(z),p1)),t) $\wedge$
HoldsAt(Touching(Robot,w,p2),t) $\wedge$
r–90 $\leq$ Bearing(p1,p2) < r+12

Happens(Switch2,t) $\leftarrow$ (H3)
Happens(Bump,t) $\wedge$ HoldsAt(Facing(r),t) $\wedge$
HoldsAt(Occupies(Robot,Displace(Disc(z),p1)),t) $\wedge$
HoldsAt(Touching(Robot,w,p2),t) $\wedge$
r–12 $\leq$ Bearing(p1,p2) < r+90

The term Occupies(Robot,Displace(Disc(z),p1)) is employed in Axioms (H2) and (H3) to obtain the centre p1 of the region occupied by the robot, which can be thought of as its *location*. Note that Axiom (H1) caters for occasions on which the robot attempts to move when it is already blocked, as well as for occasions on which the robot's motion causes it to collide with something. In the former case, an immediate Bump event occurs, and the robot accordingly moves no distance at all.

For present purposes, the Bump event is somewhat redundant. In Axioms (E5) and (E6) it could be replaced by Switch1 and Switch2 events, and in Axioms (H2) and (H3) it could be simplified away. One reason not to abolish the Bump event is that, in principle, a collision could occur without the attendant sensor event — if one of the bump switches were broken, say. Similarly, a sensor event could occur without a collision as its cause — if a rain drop were to momentarily short a connection, for example.

Another reason is that abolishing the Bump event would violate a basic principle of the present approach, according to which the assumption of an external world governed by certain physical laws, a world to which its sensors have imperfect access, is built in to the robot. The robot's task is to do its best to explain its sensor data in terms of a model

[11] Both forward bump switches are tripped if the collision point is within approximately 12° of the robot's bearing. This range was determined empirically.

of the physics governing that world. In any such model, incoming sensor data is the end of the line, causally speaking. In the physical world, it's not a sensor event that stops the robot but a collision with a solid object.

The second component of N is a description of the robot's actions. Suppose the robot behaves as illustrated in Figure 5b. Let N2 be the conjunction of the following formulae, which represent the robot's actions up to the moment when it bumps into obstacle A.

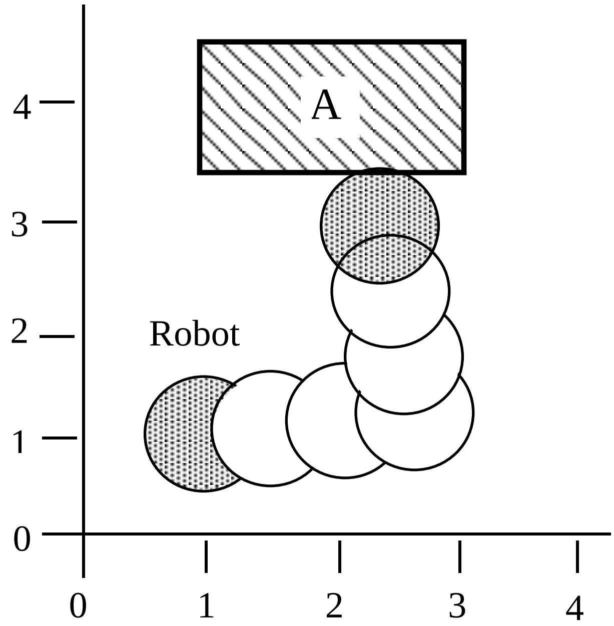

Figure 5b: A Sequence of Robot Actions

Happens(Go,0) (5.1)

Happens(Stop,2·8) (5.2)

Happens(Rotate(–90),3·3) (5.3)

Happens(Go,3·8) (5.4)

The final component of our theory is O $\wedge$ M, where M is a map of the robot's world and O is the conjunction of Axioms (Sp1) to (Sp8). Like N, M is conveniently divided into two parts. Let M be M1 $\wedge$ M2, where M1 is a description of the initial locations, shapes, and orientations (where applicable) of known objects, including the robot itself. For the example of Figure 5b, M1 would be the conjunction of the following formulae.

$\text{Initially}_{\text{P}}$(Facing(80)) (5.5)

$\text{Initially}_{\text{P}}$(Occupies(Robot,Displace(Disc(0·5),$\langle$1,1$\rangle$))) (5.6)

The form of M2 is the same as that of M1. However, when assimilating sensor data, M2 is supplied by abduction. For now though, let's look at the predictive capabilities of this framework, and supply M2 directly. Let M2 be the following formula, which describes the obstacle in Figure 5b.

$$\exists\, g\, [\mathrm{Initially_P}(\mathrm{Occupies}(A,g)) \wedge \forall\, x,y\, [\langle x,y\rangle \in g \leftrightarrow 1 < x < 3 \wedge 3{\cdot}5 < y < 4{\cdot}5]] \tag{5.7}$$

The following proposition says that, according to the formalisation, both bump switches are tripped at approximately time 5·5 (owing to a collision with obstacle A), and that the bump switches are not tripped at any other time.

Proposition 5.8.

$$\begin{aligned}&\mathrm{CIRC}\,[O \wedge M1 \wedge M2\ ;\ \mathrm{AbSpace}\ ;\ \mathrm{Initially_P}, \mathrm{Initially_N}] \wedge\\&\quad \mathrm{CIRC}[N1 \wedge N2\ ;\ \mathrm{Happens}] \wedge\\&\qquad \mathrm{CIRC}[E\ ;\ \mathrm{Initiates}, \mathrm{Terminates}, \mathrm{Releases}] \wedge B \vDash\\&\qquad\quad \mathrm{Happens}(\mathrm{Switch1},T_{bump}) \wedge \mathrm{Happens}(\mathrm{Switch2},T_{bump}) \wedge\\&\qquad\qquad [[\mathrm{Happens}(\mathrm{Switch1},t) \vee \mathrm{Happens}(\mathrm{Switch2},t)] \rightarrow t = T_{bump}]\end{aligned}$$

where $T_{bump} = \dfrac{2{\cdot}5 + 2{\cdot}8.\mathrm{Cos}(80)}{\mathrm{Cos}(-10)} + 3{\cdot}8$.

Proof. See Appendix A. □

The process of assimilating sensor data is the reverse of that of predicting sensor data. As outlined in Section 1, the task is to postulate the existence, location, and shape of a collection of objects which would explain the robot's sensor data, given its motor activity.[12]

Let Ψ be the conjunction of a set of formulae of the form Happens(Switch1,τ) or Happens(Switch2,τ) where τ is a time point. What we want to explain is the *partial completion* of this formula, for reasons that will be made clear shortly. The only-if half of this completion is defined as follows.

Definition 5.9.

$$\mathrm{COMP}[\Psi] \equiv_{def}$$

$$[\mathrm{Happens}(a,t) \wedge [a = \mathrm{Switch1} \vee a = \mathrm{Switch2}]] \rightarrow \bigvee_{\langle\alpha,\tau\rangle \in \Gamma} [a = \alpha \wedge t = \tau]$$

where $\Gamma = \{\langle\alpha,\tau\rangle \mid \mathrm{Happens}(\alpha,\tau) \in \Psi\}$. □

Given Ψ, we're interested in finding conjunctions M2 of formulae in which each conjunct has the form,

$$\exists\, g\, [\mathrm{Initially_P}(\mathrm{Occupies}(\omega,g)) \wedge \forall\, p\, [p \in g \leftrightarrow \Pi]]$$

where ω is an object constant and Π is any formula in which p is free, such that $O \wedge M1 \wedge M2$ is consistent and,

[12] In the present paper, it is assumed that all sensor data require explanation. However, to take account of glitches (as opposed to just noise), this requirement can be relaxed.

$$\begin{array}{l} \text{CIRC}[\text{O} \wedge \text{M1} \wedge \text{M2} \text{ ; AbSpace ; Initially}_\text{P}, \text{Initially}_\text{N}] \wedge \\ \quad \text{CIRC}[\text{N1} \wedge \text{N2} \text{ ; Happens}] \wedge \\ \quad\quad \text{CIRC}[\text{E ; Initiates, Terminates, Releases}] \wedge \text{B} \vDash \Psi \wedge \text{COMP}[\Psi]. \end{array}$$

The partially completed form of the Happens formula on the right-hand-side of the turnstile eliminates anomalous explanations in which, for example, the robot encounters a phantom extra obstacle before the time of the first event in Ψ. If Ψ on its own were used instead of this partially completed formula, it would be possible to construct such explanations by shifting all the obstacles that appear in a proper explanation into new positions which take account of the premature interruption in the robot's path caused by the phantom obstacle.

Clearly, from Proposition 5.8, if Ψ is,

$$\text{Happens}(\text{Switch1},\text{T}_\text{bump}) \wedge \text{Happens}(\text{Switch2},\text{T}_\text{bump})$$

then (5.7) is an explanation that meets this specification.[13] Note that the symbol A in (5.7) (or rather its computational counterpart in the actual robot), when generated through the abductive assimilation of sensor data, is *grounded* in Harnad's sense of the term [Harnad, 1990], at the same time as acquiring *meaning* through the theory. Furthermore, the theoretical framework within which such explanations are understood,

- Links the symbols that appear in them directly to a level of representation at which high-level reasoning tasks can be performed, and
- Licenses an account of the robot's success (or otherwise) at performing its tasks which appeals to the correctness of its representations and its reasoning processes.

However, (5.7) is just one among infinitely many possible explanations of this Ψ of the required form. An alternative explanation might involve the existence of an object of an entirely different shape. A bizarre example of an alternative explanation would be that the whole of space was occupied by a single object with a tunnel bored in it whose shape exactly matched that of the robot's path up to time T_bump.

In the specification of an abductive task like this, the set of explanations of the required form will be referred to as the *hypothesis space*. It's clear, in the present case, that some constraints must be imposed on the hypothesis space to eliminate bizarre explanations. Furthermore, even at a general mathematical level, the set of all explanations of the suggested form for a given stream of sensor data is hard to reason about, and computing a useful representation of such a set is infeasible. The general problem of the extravagant nature of the hypothesis space proposed above is tackled in the next section.

[13] It is assumed that the language includes an arbitrarily large set of unused constant symbols, from which ω is drawn. A similar assumption will apply throughout the sequel.

6 Boundaries

Many spatial representation techniques could be adapted for the purpose of rendering the hypothesis space more manageable (see [Davis, 1990, Chapter 6]). The one chosen here is based on the idea that a shape's boundary can be approximated to an arbitrary degree of precision by a series of straight lines.

The definition of a region's boundary is in terms of expansions of that region, defined as follows. The term Expand(g,d) denotes the region obtained by expanding g in all directions by a distance d.

$$p1 \in \text{Expand}(g,d) \leftrightarrow \exists\, p2\, [p2 \in g \wedge \text{Distance}(p1,p2) \leq d] \qquad \text{(Bo1)}$$

A set of points is a boundary of a region if it is outside that region, but any expansion of the region would engulf it. Note that this definition applies to any portion of the whole boundary, as well as to the whole boundary itself. The formula Boundary(l,g) denotes that l is a portion of the boundary of g.

$$\begin{aligned} &\text{Boundary}(g,l) \leftrightarrow \qquad \text{(Bo2)} \\ &\quad \forall\, p\, [p \in l \rightarrow [p \notin g \wedge \forall\, d\, [d > 0 \rightarrow p \in \text{Expand}(g,d)]]] \end{aligned}$$

The boundary of a shape will be described by a list of straight lines. The standard list functions Nil and Cons are introduced, and defined as follows.

$$\begin{aligned} &\text{Member}(l1,c1) \leftrightarrow \qquad \text{(Bo3)} \\ &\quad [c1 = \text{Cons}(l2,c2) \wedge [l1 = l2 \vee \text{Member}(l1,c2)]] \end{aligned}$$

$$\text{UNA}[\text{Cons, Nil}] \qquad \text{(Bo4)}$$

Every point on a boundary of the shape denoted by the term Connect(c) is also on one of the straight lines in the list c. So the term Connect(c) denotes the largest polygon enclosed by the lines in c. If the lines in c don't form a closed curve then Connect(c) is the empty set.

$$\begin{aligned} &p1 \in \text{Connect}(c) \leftrightarrow \qquad \text{(Bo5)} \\ &\quad \exists\, g\, [p1 \in g \wedge \forall\, p2,l\, [[\text{Boundary}(g,l) \wedge p2 \in l] \leftrightarrow \\ &\qquad \exists\, p3,p4\, [\text{Member}(\text{Line}(p3,p4),c) \wedge p2 \in \text{Line}(p3,p4)]] \wedge \\ &\qquad\quad \exists\, p2,z\, \forall\, p3\, [p2 \in g \rightarrow p2 \in \text{Displace}(\text{Disc}(z),p3)]] \end{aligned}$$

The last line of Axiom (Bo5) ensures that g is bounded, in the sense that a circle can be found which encloses it. This constrains g to be the inside of the shape described by c, and not the outside.

Given,

- the conjunction B of CEC with Axioms (B1) to (B6),
- the conjunction E of Axioms (E1) to (E6),
- the conjunction O of Axioms (Sp1) to (Sp8) with Axioms (Bo1) to (Bo5),

- a conjunction M1 of Initially_P and Initially_N formulae describing the initial locations, shapes, and orientations of known objects, including the robot itself,
- the conjunction N1 of Axioms (H1) to (H3),
- a conjunction N2 of Happens formulae describing the robot's actions, and
- a conjunction Ψ of formulae of the form Happens(Switch1,τ) or Happens(Switch2,τ),

we're now interested in finding conjunctions M2 of formulae in which each conjunct has the form,

$$\text{Initially}_P(\text{Occupies}(\omega,\text{Connect}(\lambda)))$$

where ω is an object constant, and λ is a term of the form,

$$\text{Cons}(\text{Line}(\rho_1,\rho_2),\text{Cons}(\text{Line}(\rho_2,\rho_3), \ldots \text{Cons}(\text{Line}(\rho_n,\rho_1),\text{Nil}) \ldots)$$

where $n \geq 3$ and $\rho_i \neq \rho_j$ for all $i \neq j$, such that $O \wedge M1 \wedge M2$ is consistent, and,

$$\begin{array}{l} \text{CIRC}[O \wedge M1 \wedge M2 \text{ ; AbSpace ; Initially}_P, \text{Initially}_N] \wedge \\ \quad \text{CIRC}[N1 \wedge N2 \text{ ; Happens}] \wedge \\ \quad\quad \text{CIRC}[E \text{ ; Initiates, Terminates, Releases}] \wedge B \vDash \Psi \wedge \text{COMP}[\Psi]. \end{array}$$

Now, returning to the example of Figure 5b, if we let M1 and N2 be the same formula as in the previous section, and let M2 be,

$$\begin{array}{l} \text{Initially}_P(\text{Occupies}(A, \\ \quad \text{Connect}(\text{Cons}(\text{Line}(\langle 1\cdot0,4\cdot5\rangle,\langle 3\cdot0,4\cdot5\rangle), \\ \quad\quad \text{Cons}(\text{Line}(\langle 3\cdot0,4\cdot5\rangle,\langle 3\cdot0,3\cdot5\rangle), \\ \quad\quad\quad \text{Cons}(\text{Line}(\langle 3\cdot0,3\cdot5\rangle,\langle 1\cdot0,3\cdot5\rangle), \\ \quad\quad\quad\quad \text{Cons}(\text{Line}(\langle 1\cdot0,3\cdot5\rangle,\langle 1\cdot0,4\cdot5\rangle))))))) \end{array} \tag{6.1}$$

then the following proposition is true.

Proposition 6.2.

$$\begin{array}{l} \text{CIRC}[O \wedge M1 \wedge M2 \text{ ; AbSpace ; Initially}_P, \text{Initially}_N] \wedge \\ \quad \text{CIRC}[N1 \wedge N2 \text{ ; Happens}] \wedge \\ \quad\quad \text{CIRC}[E \text{ ; Initiates, Terminates, Releases}] \wedge B \vDash \\ \quad\quad\quad \text{Happens}(\text{Switch1},T_{bump}) \wedge \text{Happens}(\text{Switch2},T_{bump}) \wedge \\ \quad\quad\quad\quad [[\text{Happens}(\text{Switch1},t) \vee \text{Happens}(\text{Switch2},t)] \rightarrow t = T_{bump}] \end{array}$$

where $T_{bump} = \dfrac{2\cdot5 + 2\cdot8.\text{Cos}(80)}{\text{Cos}(-10)} + 3\cdot8$.

Proof. See Appendix A. □

From Proposition 6.2, we see that (6.1) is an explanation of the sensor data which meets the above specification.

We now have a specification of the abductive task of explaining sensor data which is susceptible to both implementation and formal analysis. In addition, this theoretical framework enables us to bring formal techniques to bear on issues such as the uncertainty resulting from noise, which is addressed in the next section.

7 Noise

The hallmark of the common sense informatic situation for a mobile robot is incomplete and uncertain knowledge of a world of spatio-temporally located objects. Incompleteness is a consequence of the robot's limited window on the world, and uncertainty results from noise in its sensors and actuators. This section deals with noise.

Both noisy sensors and noisy actuators can be captured using non-determinism. (An alternative is to use probability [Bacchus, *et al.*, 1995]). Here we'll only look at the uncertainty in the robot's location that results from its noisy motors. The robot's motors are "noisy" for various reasons. For example, the two wheels might rotate at slightly different speeds when the robot is trying to travel in a straight line, or the robot might be moving on a slope or a slippery surface.[14] Motor noise of this kind can be captured using a non-deterministic Trajectory formula, such as the following replacement for Axiom (B3).[15]

$$\begin{aligned}&\exists\, p\ [\mathrm{Trajectory}(\mathrm{Moving},t,\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(g,p)),d) \wedge \\ &\quad \mathrm{Distance}(p,\langle d.\mathrm{Sin}(r),d.\mathrm{Cos}(r)\rangle) \leq d.\varepsilon] \leftarrow \\ &\qquad \mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},g),t) \wedge \mathrm{HoldsAt}(\mathrm{Facing}(r),t)\end{aligned} \tag{B7}$$

In effect, Axiom (B7) constrains the robot's location (the centre of the region it occupies) to be within an ever-expanding *circle of uncertainty* centred on the location it would be in if its motors weren't noisy.[16] The constant ε determines the rate at which this circle grows. Axiom (B8) below ensures that there are no discontinuities in the robot's trajectory. Without this axiom the robot would be able to leap over any obstacle which didn't completely cover the circle of uncertainty for its location. The term Abs(d) denotes the absolute value of d.

$$\begin{aligned}&\mathrm{Trajectory}(f,t,\mathrm{Occupies}(x,\mathrm{Displace}(g,p1)),d1) \rightarrow \\ &\quad \forall\, z\ [z > 0 \rightarrow \exists\, d\ \forall\, d2,p2\ \ [[d2 > 0 \wedge \mathrm{Abs}(d2-d1) < d \wedge \\ &\qquad \mathrm{Trajectory}(f,t,\mathrm{Occupies}(x,\mathrm{Displace}(g,p2)),d2)] \rightarrow \\ &\qquad\quad \mathrm{Distance}(p1,p2) < z]]\end{aligned} \tag{B8}$$

Figure 7a shows the robot exploring the corner of an obstacle. Figure 7b shows the evolution of the corresponding circle of uncertainty, highlighting the points where the robot changes direction.

---

[14] The Rug Warrior's wheels are fitted with shaft encoders, which can be used to reduce the non-determinism in the robot's movements. These are ignored in the present analysis.

[15] The Rotate action could also be made non-deterministic, yielding only a range of possible values for the robot's subsequent orientation.

[16] Note that, while objects occupy open subsets of $\mathbb{R}^2$, regions of uncertainty are closed.

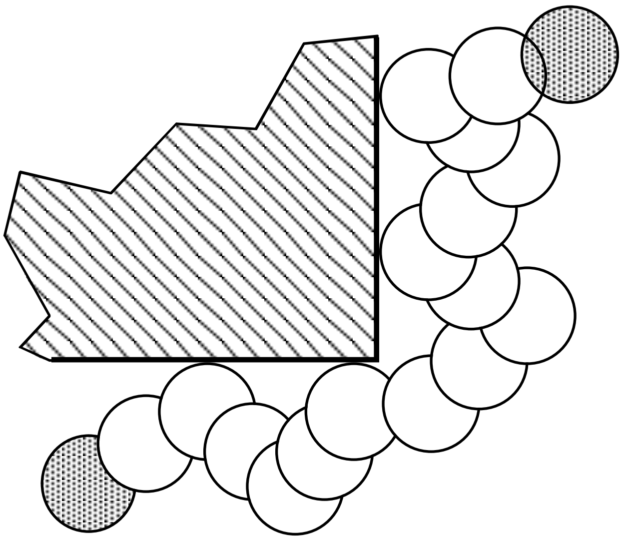

Figure 7a: The Robot Explores a Corner

Figure 7b is somewhat misleading, however. Consider Figure 7c. On the left, the evolution of the circle of uncertainty for the robot's location is shown. In the middle, three potential locations are shown for the three changes of direction. Although these locations all fall within the relevant circles of uncertainty, the robot could never get to the third location from the second. This is because, as depicted on the right of the figure, in any given model the circle of uncertainty for the robot's location at the end of a period of continuous motion can only be defined relative to its actual location at the start of that period. This can be verified by inspecting Axioms (B7) and (B8).

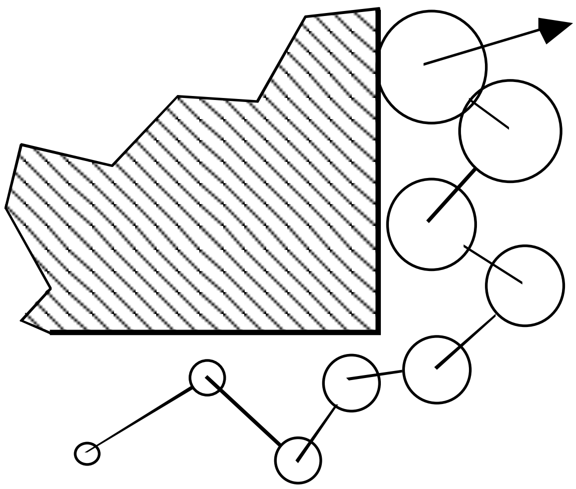

Figure 7b: The Evolution of the Circle of Uncertainty

The relative nature of the evolution of the circle of uncertainty means that the robot can acquire a detailed knowledge of some area A1 of its environment, then move to another area A2 which is some distance from A1, and acquire an equally detailed knowledge of A2. The accumulated uncertainty entails only that the robot is uncertain of where A1 is relative to A2. This natural feature of the formalisation conforms with what we would intuitively expect given the robot's informatic situation.

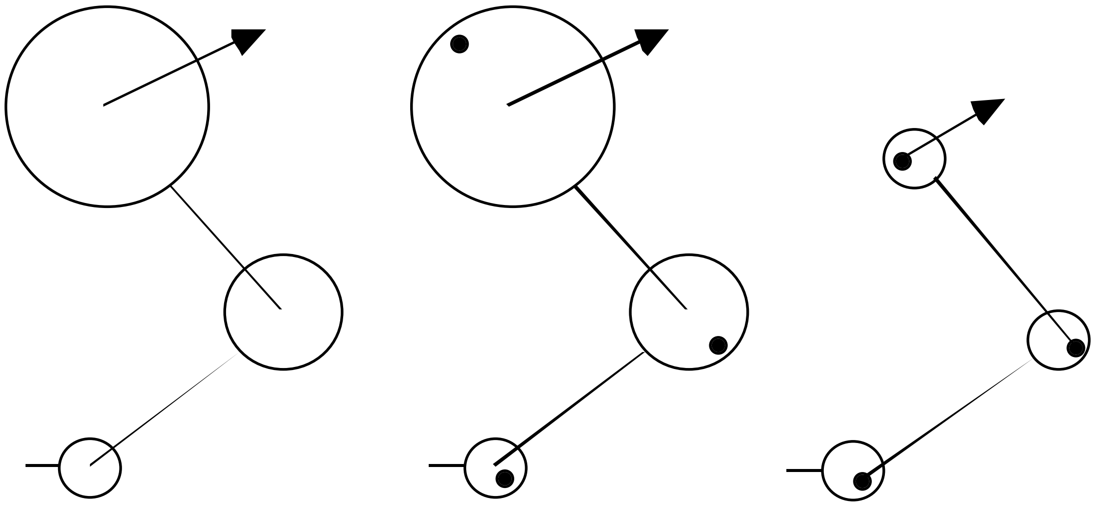

Figure 7c: The Circle of Uncertainty is Relative Not Absolute

Non-determinism is a potential source of difficulty for the abductive approach to explanation. Even with a precise and complete description of the initial state of the world, including all its objects and their shapes, a non-deterministic theory incorporating a formula like (B7) will not yield the exact times at which collision events occur. Yet the sensor data to be assimilated has precise times attached to it. Fortunately we can recast the task of assimilating sensor data as a form of *weak abduction* so that it yields the required results. Intuitively what we want to capture is the fact that without the hypothesised objects, the sensor data could not have been received. This is analogous to the consistency-based approach to diagnosis proposed by Reiter [1987].

Definition 7.1. Given,

- the conjunction B of CEC with Axioms (B1), (B2), and (B4) to (B8),
- the conjunction E of Axioms (E1) to (E6),
- the conjunction O of Axioms (Sp1) to (Sp8) with Axioms (Bo1) to (Bo5),
- a conjunction M1 of Initially_P and Initially_N formulae describing the initial locations, shapes, and orientations of known objects, including the robot itself,
- the conjunction N1 of Axioms (H1) to (H3),
- a conjunction N2 of Happens formulae describing the robot's actions, and
- a conjunction Ψ of formulae of the form Happens(Switch1,τ) or Happens(Switch2,τ),

an *explanation* of Ψ is a conjunction M2 of formulae in which each conjunct has the form,

$\text{Initially}_P(\text{Occupies}(\omega,\text{Connect}(\lambda)))$

where ω is an object constant, and λ is a term of the form,

$\text{Cons}(\text{Line}(\rho_1,\rho_2),\text{Cons}(\text{Line}(\rho_2,\rho_3), \ldots \text{Cons}(\text{Line}(\rho_n,\rho_1),\text{Nil}) \ldots))$

where $n \geq 3$ and $\rho_i \neq \rho_j$ for all $i \neq j$, such that $O \wedge M1 \wedge M2$ is consistent, and,

$$\text{CIRC}[O \wedge M1 \wedge M2 \,;\, \text{AbSpace} \,;\, \text{Initially}_P, \text{Initially}_N] \wedge$$
$$\text{CIRC}[N1 \wedge N2 \,;\, \text{Happens}] \wedge$$
$$\text{CIRC}[E \,;\, \text{Initiates, Terminates, Releases}] \wedge B \not\models \neg [\Psi \wedge \text{COMP}[\Psi]]. \quad \square$$

To illustrate the new definition, suppose the robot behaves as illustrated in Figure 7d. Let N2 be the conjunction of the following formulae, which represent the robot's actions up to and including the time it bumps into obstacle A.

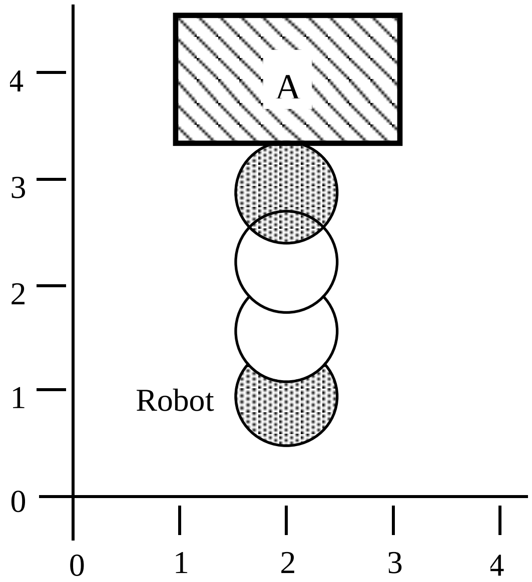

Figure 7d: The Robot Collides with an Obstacle

$\text{Happens}(\text{Go},0)$ (7.2)

$\text{Happens}(\text{Stop},2\cdot 1)$ (7.3)

Let M1 be the conjunction of the following formulae.

$\text{Initially}_P(\text{Facing}(0))$ (7.4)

$\text{Initially}_P(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0\cdot 5),\langle 2,1\rangle)))$ (7.5)

Let M2 be the following formula.

$$\begin{array}{l}\text{Initially}_P(\text{Occupies}(A, \\ \quad \text{Connect}(\text{Cons}(\text{Line}(\langle 1\cdot 0,4\cdot 5\rangle,\langle 3\cdot 0,4\cdot 5\rangle), \\ \quad\quad \text{Cons}(\text{Line}(\langle 3\cdot 0,4\cdot 5\rangle,\langle 3\cdot 0,3\cdot 5\rangle), \\ \quad\quad\quad \text{Cons}(\text{Line}(\langle 3\cdot 0,3\cdot 5\rangle,\langle 1\cdot 0,3\cdot 5\rangle), \\ \quad\quad\quad\quad \text{Cons}(\text{Line}(\langle 1\cdot 0,3\cdot 5\rangle,\langle 1\cdot 0,4\cdot 5\rangle)))))))\end{array} \quad (7.6)$$

In the noise-free case, the robot would collide with A at time 2·0. However, let's assume the collision takes place at time 2·1. Let Ψ be the conjunction of the following formulae.

$$\text{Happens(Switch1,2·1)} \tag{7.7}$$

$$\text{Happens(Switch2,2·1)} \tag{7.8}$$

Let ε be 0·25. The following proposition says that M2 is indeed an explanation of Ψ according to the new definition.

Proposition 7.9.

$$\begin{aligned}&\text{CIRC [O} \wedge \text{M1} \wedge \text{M2 ; AbSpace ; Initially}_\text{P}\text{, Initially}_\text{N}] \wedge \\ &\quad \text{CIRC[N1} \wedge \text{N2 ; Happens]} \wedge \\ &\quad\quad \text{CIRC[E ; Initiates, Terminates, Releases]} \wedge \text{B} \nvDash \neg [\Psi \wedge \text{COMP}[\Psi]].\end{aligned}$$

Proof. See Appendix A. □

There will, naturally, be many explanations for any given Ψ which meet Definition 7.1. For the example above, an explanation which postulated an object of the same shape but 0·1 units of distance further north would have sufficed. A standard way to treat multiple explanations in abductive knowledge assimilation is to adopt their disjunction [Shanahan, 1997a, Chapter 17]. This has the effect of smothering any explanations which are stronger than necessary, such as those which postulate superfluous obstacles. The disjunction of all explanations of Ψ is the *cautious explanation* of Ψ.

A variety of *preference relations* over explanations can also be introduced, which make the abductive process more selective. For example, it might be reasonable to assume that nearby collision points indicate the presence of a single object. This is captured in the following definitions.

Definition 7.10. Two objects are *separated by* x units if there is a line of length x which has points in common with the boundaries of both objects, and there is no shorter such line. □

Definition 7.11. An explanation M_1 is *preferable* to an explanation M_2 if the number of objects in M_1 which are separated from another object by less than MIN units is less than the number of such objects in M_2. □

A *preferred* explanation is one for which no preferable explanation exists. Such preference relations are a topic for further study, and will not feature in the rest of the paper.

The following theorem establishes that the above definition of an explanation is equivalent to the deterministic specification offered in the last section when ε is 0. Let B_{det} be the conjunction of CEC with Axioms (B1) to (B6).

Definition 7.12. A formula M is a *complete spatial description* if the region occupied by each object mentioned in M is the same in every model of,

$\text{CIRC}[\text{O} \wedge \text{M} ; \text{AbSpace} ; \text{Initially}_\text{P}, \text{Initially}_\text{N}]$. □

Theorem 7.13. If $\varepsilon = 0$ and M1 is a complete spatial description, then M2 is an explanation of Ψ if and only if O $\wedge$ M1 $\wedge$ M2 is consistent and,

$$\begin{aligned}&\text{CIRC}[\text{O} \wedge \text{M1} \wedge \text{M2} ; \text{AbSpace} ; \text{Initially}_\text{P}, \text{Initially}_\text{N}] \wedge \\ &\quad \text{CIRC}[\text{N1} \wedge \text{N2} ; \text{Happens}] \wedge \\ &\quad\quad \text{CIRC}[\text{E} ; \text{Initiates, Terminates, Releases}] \wedge \text{B}_\text{det} \vDash \Psi \wedge \text{COMP}[\Psi].\end{aligned}$$

Proof. See Appendix B. □

Note that, as illustrated in Figure 7e, while the location of the robot at any time falls within a circle of uncertainty, the corresponding region of uncertainty for a collision point is kidney-shaped, since the collision point could be anywhere along a portion of the circumference of the robot. (Recall that the forward bump switches deliver three possible values for the direction of a collision.)

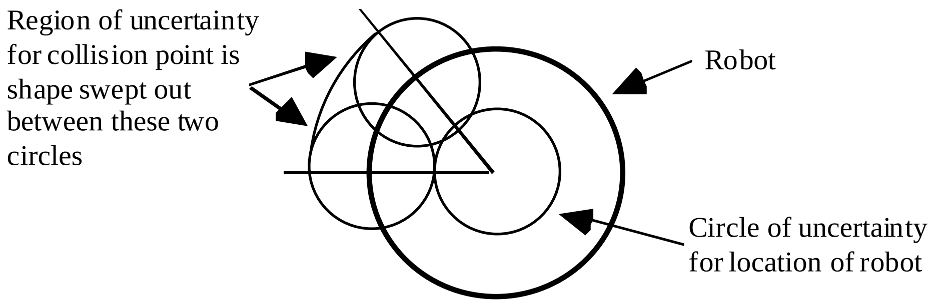

Figure 7e: The Region of Uncertainty for a Collision Point

In the next section, two results are established which support the claim that the abductive specification of the task of sensor data assimilation offered here gives the results we would expect, and which also make the task of building algorithms based on this specification straightforward.

8 Two Theorems

The two theorems in this section are an attempt to characterise the abductive explanations defined in the previous section in terms which appeal more directly to the information available to any map-building algorithm which might be executed on board the robot. This is done by picking out the regions of uncertainty within which the collision points corresponding to a stream of switch events must lie.

Let B be the conjunction of CEC with Axioms (B1), (B2), and (B4) to (B8). Let E be the conjunction of axioms (E1) to (E6). Let O be the conjunction of Axioms (Sp1) to (Sp8) with Axioms (Bo1) to (Bo5). Let M1 be a conjunction of $\text{Initially}_\text{P}$ and $\text{Initially}_\text{N}$ formulae describing the initial locations, shapes, and orientations of known objects, including the robot itself. Let N1 be the conjunction of Axioms (H1) to (H3).

Let N2 be a conjunction of Happens formulae describing the robot's actions which is equivalent to the formula,

$\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_n$

where each ϕ_i has the form Happens(α_i,τ_i), and,

- for all $i < n$, $\tau_i \leq \tau_{i+1}$ if α_i is Go and α_{i+1} is Stop,
- for all $i < n$, $\tau_i < \tau_{i+1}$ if α_i is not Go or α_{i+1} is not Stop,
- α_1 is Go, and
- for all $i < n$, α_{i+1} is $\begin{cases} \text{Stop if } \alpha_i \text{ is Go,} \\ \text{Rotate}(\theta) \text{ if } \alpha_i \text{ is Stop,} \\ \text{Go if } \alpha_i \text{ is Rotate}(\theta). \end{cases}$

N2 describes a finite sequence of actions of the form,

Go, Stop, Rotate, Go, Stop, Rotate, Go, . . .

The i^{th} action in N2 occurs before the $i+1^{th}$ action, unless the i^{th} action is a Go and the $i+1^{th}$ is a Stop, in which case they can occur simultaneously. (A Go and a Stop will occur simultaneously if the robot tries to move when it is already pressed up against an obstacle.) Let Ψ be a conjunction of the form,

$\psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_m$

where each ψ_i is either a *type 1 datum* of the form,

Happens(Switch1,τ_i)

or a *type 2 datum* of the form,

Happens(Switch2,τ_i)

or a *type 3 datum* of the form,

Happens(Switch1,τ_i) $\wedge$ Happens(Switch2,τ_i)

such that for all $i < m$, $\tau_i < \tau_{i+1}$.

Definition 8.1. N2 *tracks* Ψ if for every occurrence in Ψ of a formula of the form,

Happens(α,τ)

where α is either Switch1 or Switch2, there is a corresponding occurrence in N2 of the formula,

Happens(Stop,τ). □

We will assume that N2 tracks Ψ.[17] This is the case if the robot's control mechanism ensures that it stops moving immediately when it bumps into something. This makes the theorems below much easier to state and prove, since it enables us to ignore the possibility of the robot's wheels spinning while it is pressed up against an obstacle.

[17] In practice, of course, there's always a small delay between the Bump event and the Stop event. This delay is so small it can safely be ignored.

Let $T_{go}(i)$ and $T_{stop}(i)$ be respectively the times of the i^{th} Go action in N2 and the i^{th} Stop action in N2. Where no such action exists, these functions are undefined.

The MT function, defined below, yields the amount of time the robot has been in motion between the j^{th} and the k^{th} Stop actions. The radius of the circle of uncertainty for the robot's location at the time of the k^{th} Stop action relative to its location at the time of the j^{th} Stop action is proportional to this value.

Definition 8.2. The function MT (for Motion Time) is defined as follows.

$$MT(j,k) = \sum_{j<i\leq k} (T_{stop}(i) - T_{go}(i)). \qquad \square$$

Note that MT(0,k) is the amount of time the robot has been in motion up to the time of the k^{th} Stop action.

The RO function, defined below, yields the orientation of the robot at a given time.

Definition 8.3. The function RO (for Robot Orientation) is defined as follows.

$$RO(\tau) = \theta_0 + \sum_{i\leq R} (\theta_i)$$

where R is the number of Rotate actions in N2 before τ, the i^{th} Rotate action in N2 is Happens(Rotate(θ_i),τ_i), and θ_0 is the robot's initial orientation according to M1. $\square$

The RL function, defined below, yields the centre of the circle of uncertainty for the robot's location at the time of the k^{th} Stop action given that its location is ρ at the time of the j^{th} Stop action.

Definition 8.4. The function RL (for Robot Location) is defined as follows.

$$RL(\rho,j,k) = \rho + \sum_{j<i\leq k} \rho_i$$

where,

- ρ_i is $\langle \delta_i.Sin(\theta_i),\delta_i.Cos(\theta_i)\rangle$,
- δ_i is $T_{stop}(i) - T_{go}(i)$, and
- θ_i is $RO(T_{go}(i))$. $\square$

Note that, if ρ is the robot's initial location, RL(ρ,0,k) is the centre of the circle of uncertainty for the robot's location at the time of the k^{th} Stop action.

The BD function, defined below, yields the direction of the centre of the region of uncertainty for a collision (see Figure 7d).

Definition 8.5. The functions BD_{min} and BD_{max} (for minimum and maximum Bump Direction) are defined as follows.

$$BD_{min}(n,\theta) = \begin{cases} \theta - 90 \text{ if } n = 1 \\ \theta + 12 \text{ if } n = 2 \\ \theta - 12 \text{ if } n = 3 \end{cases}$$

$$BD_{max}(n,\theta) = \begin{cases} \theta - 12 \text{ if } n = 1 \\ \theta + 90 \text{ if } n = 2 \\ \theta + 12 \text{ if } n = 3 \end{cases}$$ □

Let η_i be the type of datum ψ_i and let τ_i be the time of that datum. A bump trail, defined formally below, is a sequence of points, one for each collision featured in Ψ. The i^{th} point in this sequence ρ_i is a possible location, according to N2 and the axioms constraining the robot's motion, for the robot at the time of the i^{th} collision, given that the robot was at ρ_{i-1} at the time of the $i\text{–}1^{th}$ collision (see Figure 7d).

Definition 8.6. A sequence of points ρ_1 to ρ_m is a *bump trail* if,

- the distance from ρ_1 to the point RL(ρ_0,0,k) is less than or equal to ε.MT(0,k), where T_{stop}(k) is τ_1 and ρ_0 is the robot's initial location according to M1, and
- for any i, $0 < i < m$, the distance from ρ_{i+1} to the point RL(ρ_i,j,k) is less than or equal to ε.MT(j,k), where T_{stop}(j) is τ_i and T_{stop}(k) is τ_{i+1}. □

Now we define the relation OTR, which is used to state two theorems which characterise the set of explanations of a given stream of sensor data directly in terms of the robot's actions, and which justify the map building algorithm presented in the next section.

Definition 8.7. The relation OTR (for Obstacle Touches Robot) is defined as follows.

$$\begin{aligned} &\text{OTR(p1,r1,r2)} \equiv_{def} \\ &\quad \exists \text{ p2 [r1} \leq \text{Bearing(p1,p2)} < \text{r2} \wedge \text{Distance(p1,p2)} = 0{\cdot}5 \wedge \\ &\quad\quad \exists \text{ w,g [Initially}_P\text{(Occupies(w,g))} \wedge \text{w} \neq \text{Robot} \wedge \\ &\quad\quad\quad \text{p2} \in \text{Boundary(g)} \wedge \neg \exists \text{ p [p} \in \text{Line(p1,p2)} \wedge \text{p} \in \text{g]]]} \end{aligned}$$ □

Less formally, OTR(p1,r1,r2) is true if there is a line from p1 to p2 such that,

- p2 is at a bearing of between r1 and r2 degrees from p1 and 0·5 units away from p1,
- p2 is on the boundary of an object, and
- no point on the line from p1 to p2 is inside that object.

OTR(p,r1,r2) is true if there is an object with a boundary which would touch the robot's circumference at a suitable point if the robot were located at p. The last of the above conditions ensures that the robot touches the outside of the object rather than the inside. The first theorem using the OTR relation sets out a necessary condition for M2 to be an explanation.

Theorem 8.8. If M2 is an explanation of Ψ then there exists a bump trail ρ_1 to ρ_m such that, for any i, $0 < i \leq m$,

$$\text{O} \wedge \text{M1} \wedge \text{M2} \vDash \text{OTR}(\rho_i,BD_{min}(\eta_i,RO(\tau_i)),BD_{max}(\eta_i,RO(\tau_i))).$$

Proof. See Appendix B. □

The second theorem partially defines the sufficient conditions for M2 to be an explanation. Only a partial characterisation is possible using OTR alone. A complete characterisation would have to take account of the constraints imposed on the spatial occupancy of hypothesised objects by the path along which the robot has moved without bumping into anything.

Theorem 8.9. M2 is an explanation of Ψ if O $\wedge$ M1 $\wedge$ M2 is consistent, there exists a bump trail ρ_1 to ρ_m such that, for any i, $0 < i \leq m$,

$$O \wedge M1 \wedge M2 \models OTR(\rho_i, BD_{min}(\eta_i, RO(\tau_i)), BD_{max}(\eta_i, RO(\tau_i)))$$

and,

$$\begin{aligned}&CIRC[O \wedge M1 \wedge M2 \ ; AbSpace \ ; Initially_P, Initially_N] \wedge \\ &\quad CIRC[N1 \wedge N2 \ ; Happens] \wedge \\ &\qquad CIRC[E \ ; Initiates, Terminates, Releases] \wedge B \not\models \neg\, COMP[\Psi].\end{aligned}$$

Proof. See Appendix B. □

A similar proof to that of Theorem 8.9 establishes the following consistency result.

Theorem 8.10. If O $\wedge$ M1 $\wedge$ M2 is consistent then so is,

$$\begin{aligned}&CIRC[O \wedge M1 \wedge M2 \ ; AbSpace \ ; Initially_P, Initially_N] \wedge \\ &\quad CIRC[N1 \wedge N2 \ ; Happens] \wedge \\ &\qquad CIRC[E \ ; Initiates, Terminates, Releases] \wedge B.\end{aligned}$$

Proof. See Appendix B. □

9 An Algorithm for Sensor Data Assimilation

One approach to building an implementation based on the the logical framework presented in this paper would be to isolate a fragment of the theory which could be executed as a logic program. A body of techniques for abductive logic programming has been developed, and these could be used to implement the abductive approach to sensor data assimilation [Kakas, Kowalski & Toni, 1993]. Taking a logic programming approach would make clear the sense in which the logical account given above supplies the meaning of the representations employed by the robot.

An alternative approach is to tailor-make algorithms for specific tasks, such as sensor data assimilation, whose correctness with respect to the logical account can be demonstrated. This is the methodology I will adopt here, and the logic programming approach is left for further research.

In both approaches, the products of the abductive process should be considered as general purpose knowledge with a declarative meaning. Theorem proving techniques, such as those employed in logic programming, are still applicable to the subsequent processing of this knowledge, in order to construct plans or to answer queries, for example. A tailor-made algorithm for sensor data assimilation, in so far as it can be validated in terms of the theoretical framework supplied, can be thought of as an efficient but highly specialised theorem prover.

The two theorems in the previous section can be used to justify a variety of algorithms for constructing a map of the robot's environment given a stream of sensor data. The algorithm described in this section works with a grid of squares superimposed on $\mathbb{R}^+ \times \mathbb{R}^+$.[18] Each square is denoted by a pair of co-ordinates, each of which is a natural number. This grid is represented in a two-dimensional *occupancy array* M (see [Davis, 1990, Section 6.2.1]). If M[x,y] = c, where c is either Black, White or Grey, then the square $\langle x,y \rangle$ is coloured c. Here is the main algorithm.

Procedure ColourMap
Let every element of M be Grey
P[0] := the robot's initial location according to M1
j := 0
For i = 1 to m
    Find k such that the k^{th} Stop action in N2 occurs at τ_i
    PlotPath(P[i–1],q,j,k)
    p := RL(q,k–1,k)
    e := ε.MT(k–1,k)
    b1 := $BD_{min}(\eta_i, RO(\tau_i))$; b2 := $BD_{max}(\eta_i, RO(\tau_i))$
    Choose P[i], x, and y such that PossSquare(P[i],p,b1,b2,e,x,y)
    ColourPath(q,P[i])
    If M[x,y] = White
    Then Exit
    Else M[x,y] := Black
    j := k
End For
Output M

This non-deterministic procedure colours every square either black, white or grey, where black represents that the corresponding square region of $\mathbb{R}^+ \times \mathbb{R}^+$ is occupied by some object and white represents that it is empty. A boundary between two black squares is considered occupied while a boundary between a black and a white square is considered unoccupied, although in the algorithm description, a square region is assumed to include its boundaries (so a boundary is included in two squares and a corner is included in four). Let μ be the the length of the sides of a square.

Definition 9.1. The point $\langle a,b \rangle$ is *included* in the square $\langle x,y \rangle$ if and only if $x.\mu \leq a \leq (x+1).\mu$ and $y.\mu \leq b \leq (y+1).\mu$. □

[18] Without loss of generality, we can work in one quadrant of the plane.

Let Ψ, N2, and M1 be defined as in the previous section, along with τ_i, η_i, and the functions RL, MT, BD_{min} and BD_{max}.

Some explanation of the algorithm is in order. To begin with all squares are coloured grey. The outer loop works through each of the m collisions in Ψ in the order in which they occurred, assigning a location to the robot for each one. This sequence of locations corresponds to a bump trail, as defined in the previous section, and is recorded in the array P. For each collision, a suitable path is first plotted which takes the robot up to the site of its last turn before that collision (Line 7).

The procedure PlotPath is defined as follows. This procedure non-deterministically plots a path from point p1, which is robot's location at the time of the j^{th} Stop action, to p2, which is its location at the time of the $k–1^{th}$ Stop action. The intermediate Stop actions are those which don't correspond to collisions.

1. **Procedure** PlotPath(p1,p2,j,k)
2. **For** l = j+1 to k–1
3.   p := RL(p1,l–1,l)
4.   e := ε.MT(l–1,l)
5.   Choose any p2 whose distance from p is less than or equal to ε.MT(j,k)
6.   ColourPath(p1,p2)
7.   p1 := p2
8. **End** For

The path computed by PlotPath cuts through a region of space which is big enough for the robot to pass through. This region has to be coloured white, so that no attempt is made to place object inside it. This is the purpose of the procedure ColourPath, which is defined as follows.

1. **Procedure** ColourPath(p1,p2)
2. **For** each x and y such that Covered(x,y,p1,p2)
3.   **If** M[x,y] = black
4.   **Then Exit**
5.   **Else** M[x,y] := white
6. **End** For

Covered(x,y,p1,p2) is true if and only if the square ⟨x,y⟩ includes a point less than ζ units of distance from any point on the straight line from p1 to p2, where ζ is the radius of the robot.

Returning to the main procedure, the robot has now been taken up to the site of its last change of direction before a collision. Next, a point is chosen for its location at the time of the collision such that the robot will abut a square (Line 11). This square can then be coloured black. The boolean function PossSquare(p1,p2,b1,b2,e,x,y), which picks out candidate squares, returns True if,

- the distance from p1 to p2 is less than or equal to e,
- a straight line emanating from p1 at a bearing b, where $b1 \leq b < b2$ crosses a boundary and enters the square $\langle x,y \rangle$ at a distance of 0·5 units, and
- $M[x,y] \neq$ White

and returns False otherwise.

If ColourMap terminates with output M, any squares in M that are still grey can be coloured either black or white, possibly to conform with some preference relation over explanations such as the example in Section 7. If ColourMap exits via Line 14 or via Line 4 of ColourPath, then its non-deterministic choices have not led to a consistent colouring of M. By augmenting the algorithm with a suitable search strategy, alternative choices can be explored.[19]

The following theorems relate this algorithm to the abductive characterisation of sensor data assimilation offered in the preceding sections. First we need to be able to convert the output of ColourMap to a logical form.

Definition 9.2. If M is an occupancy array, let Exp*(M) denote the set of all conjunctions of formulae in which each conjunct has the form,

$$\text{Initially}_P(\text{Occupies}(\omega,\text{Connect}(\lambda)))$$

where ω is an object constant, and λ is a term of the form,

$$\text{Cons}(\text{Line}(\rho_1,\rho_2),\text{Cons}(\text{Line}(\rho_2,\rho_3),\ ...\ \text{Cons}(\text{Line}(\rho_n,\rho_1),\text{Nil})\ ...\ ))$$

where $n \geq 3$ and $\rho_i \neq \rho_j$ for all $i \neq j$, such that the term $\text{Line}(\rho_i,\rho_{i+1})$ occurs in Exp*(M) if and only if the line from ρ_i and ρ_{i+1} is a boundary shared by one black square and one grey or white square according to M. □

Clearly the members of Exp*(M) are all logically equivalent, modulo changes in the names of objects. Accordingly, let's suppose the existence of some standard ordering over the members of any given Exp*(M), and let Exp(M) denote the first in this ordering.

As usual, let B be the conjunction of CEC with Axioms (B1), (B2), and (B4) to (B8). Let E be the conjunction of axioms (E1) to (E6). Let O be the conjunction of Axioms (Sp1) to (Sp8) with Axioms (Bo1) to (Bo5). And let N1 be the conjunction of Axioms (H1) to (H3).

The following theorem expresses the soundness of the algorithm.

Theorem 9.3. If ColourMap terminates with output M then Exp(M) is an explanation of Ψ.

---

[19] Although each non-deterministic choice of a location in the algorithm has uncountably many possibilities, there are only finitely many squares that each location can fall within. So an algorithm that explores the whole search space in finite time may be possible.

Proof. See Appendix B. □

The ColourPath procedure above assumes that the robot always travels in a straight line (although the bearing of that line and how far the robot travels in a given time is uncertain). The algorithm misses explanations which depend crucially on the possibility, allowed by Axioms (B7) and (B8), of the robot curving around an obstacle to avoid it. The following formula, which is a strengthening of Axiom (B8), reflects this limitation, and permits us to state a theorem expressing the extent to which the algorithm is complete.

$$\exists \text{ p2,p3 } \forall \text{ d [Trajectory(f,t,Occupies(x,Displace(g,p1)),d)} \rightarrow \text{p1} \in \text{Line(p2,p3)]} \quad \text{(B9)}$$

Definition 9.4. An explanation M2 of Ψ is *curve-free* if,

$$\text{CIRC[O} \wedge \text{M1} \wedge \text{M2 ; AbSpace ; Initially}_\text{P}\text{, Initially}_\text{N}] \wedge \text{CIRC[N1} \wedge \text{N2 ; Happens]} \wedge \text{CIRC[E ; Initiates, Terminates, Releases]} \wedge \text{B} \wedge \text{(B9)} \nvDash \neg \text{COMP}[\Psi]$$ □

Theorem 9.5. For every occupancy array M such that Exp(M) is a curve-free explanation of Ψ, M is output by some execution of ColourMap.

Proof. See Appendix B. □

The final question, to be briefly considered here, is which explanations are excluded by the demands on the form of an explanation made by Theorem 9.5. Clearly, for a given square-size μ, explanations for Ψ can exist which are not of the form Exp(M) for any M. Any explanation which postulated the existence of a triangular object would be an example. On the other hand, it's also clear that there's a sense in which, by making μ small enough, any object's shape can be approximated to any desired degree of accuracy.

But the value of μ isn't the only consideration here. The radius of the robot and the value of ε, the uncertainty factor in the robot's location, are also crucial. To see this, suppose ε is zero, and let the radius of the robot be $\frac{\mu}{10}$. Now suppose the robot collides with an obstacle, rotates 180° moves forward $\frac{\mu}{2}$ units and collides with another obstacle. Clearly no explanation of the required form exists for these collisions, because no suitable pair of squares can exist with boundaries sufficiently close together. Given the radius of the robot, even the uncertainty in the whereabouts on the robot's circumference of the collision point doesn't permit a suitable square to be found. An appropriate choice of μ for a given ε should render such cases pathological, but the mathematical basis for this claim remains to be worked out.

10 Experiments with the Robot

This section reports the results of some preliminary experimentation with the Rug Warrior robot based on the algorithm described in the preceding section. The implementation is rather crude, and should be regarded as nothing more than a demonstration of the feasibility of robotics research in the advocated style.

Control of the robot was achieved with a simple two-layered Brooks-style behaviour-based architecture [Brooks, 1986]. A Wander behaviour and an Avoid behaviour were implemented. The Wander behaviour executes a random walk around the environment. When this behaviour is active, the robot moves forwards for a certain distance, subject to random variation, then rotates a certain amount, also subject to random variation, then moves forwards again, and so on.

The Avoid behaviour is triggered when one or more of the forward bump switches is tripped. When the Avoid behaviour is active, after backing a small distance away from the (presumed) obstacle, the robot rotates a certain amount, subject again to random variation. Control is then relinquished, and the Wander behaviour becomes active once more.

In the background, a process is run continuously which records the robot's actions and the sensor events that occur. This data is subsequently processed by a program which incorporates the sensor data assimilation algorithm proposed in the previous section, and which generates an occupancy array. In the present implementation, this processing is done off-board, although this purely for convenience in program development and debugging. Both components — the behaviour-based control module and the sensor data assimilation module — are written in C, the former in a dialect which includes multi-tasking features which facilitate the construction of behaviour-based control systems. The control component comprises about 300 lines of code, and the assimilation program about 900 lines.

The present implementation of the sensor data assimilation algorithm performs a negligible amount of search, and is accordingly not always successful at explaining a given set of sensor data. It always chooses the centre of the circle of uncertainty for the robot's location when finding a path from one collision site to the next. Only when choosing the robot's location at the time of a collision does the algorithm exploit the existence of the circle of uncertainty by nudging the robot's location so that it abuts a square. The point on the robot's circumference which touches this square depends on whether the collision in question involved just Switch 1, just Switch 2, or both switches. In the first two cases, the chosen point is respectively –45° and 45° from the robot's bearing. In the third case, the chosen point is directly ahead of the robot in the direction in which it is facing.

A small amount of search is carried out to determine the choice of abutting squares. If, because of occupancy data already present in the array, a location for the robot cannot be found which is within the square in which the centre of the circle of uncertainty for its location lies, then the surrounding four squares — north, east, south, and west — are also tried.

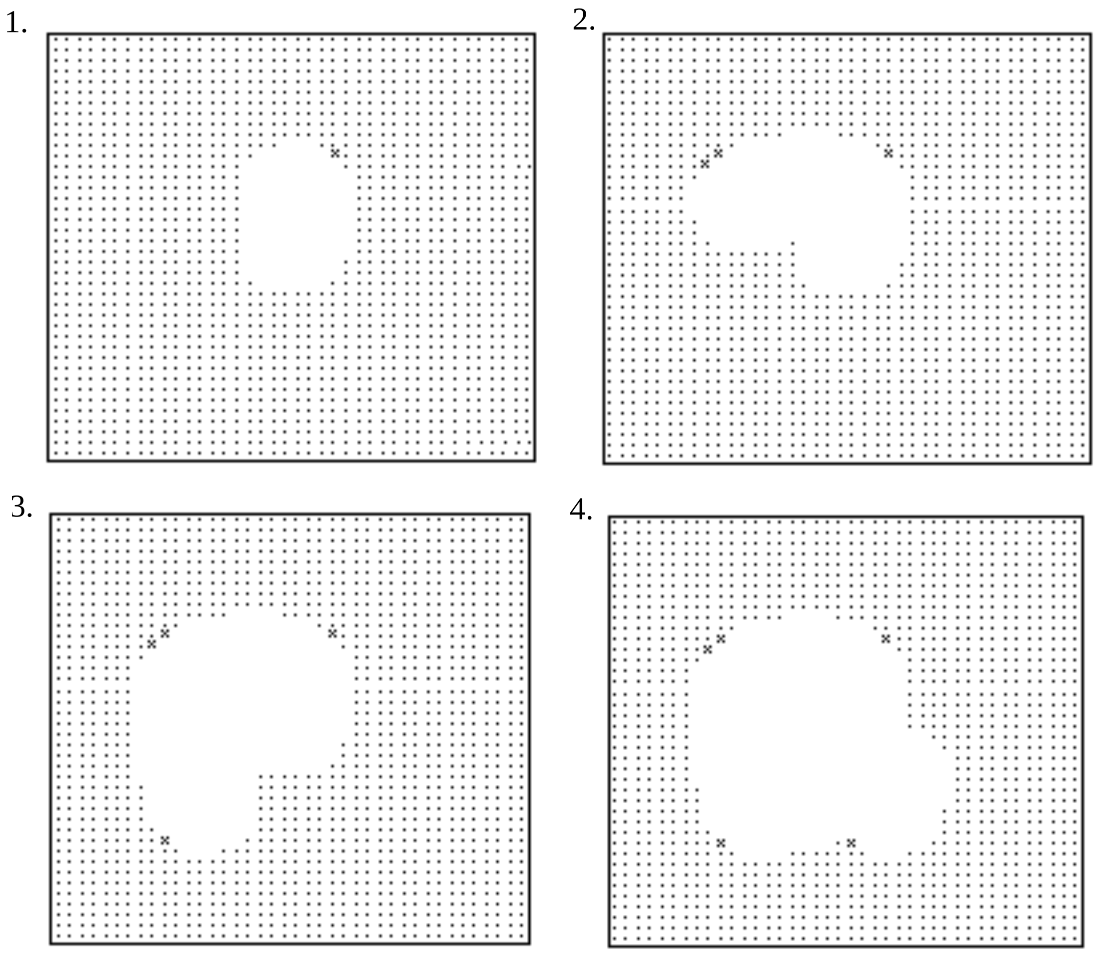

Figure 10a: The Evolution of the Occupancy Array

Figure 10a shows the evolution of the occupancy array generated by a sample run of the robot as it moving about inside a pen made up from four box files. The array is 40 squares by 40 squares, and each square is 0·25 units of distance on each side. (Recall that one robot radius is the unit of distance.) The value of ε was 1·0. The dots denote grey squares (neither occupied nor unoccupied), the crosses denote black squares, and the blank area comprises white squares. The robot's actual environment is depicted in Figure 10b, which is a scale drawing traced from an aerial photograph. As Figure 10a illustrates dramatically, with only a few collisions to go on, the robot will acquire more knowledge about empty space than about occupied space. However, given the severe limitations of the robot's perceptual apparatus, this is unsurprising.

The sample run shown is rather unimpressive since, although thirty actions were performed, only five Bump events took place. By allowing the implementation to perform more search, it is anticipated that much longer runs will be feasible, and more detailed maps will be produced. But the issue of future improvements in the implementation, though important, is a distraction from the real point here, which is that the occupancy array depicted has a logical interpretation according to which it is an explanation, in the formal sense defined in the foregoing sections, of the robot's sensor data given its actions. The benefits of this are twofold. First, the relationship between the robot's model of the world and the world itself is precisely defined. Second, a wealth of techniques and results from the field of Knowledge Representation is made available for the robot's further development.

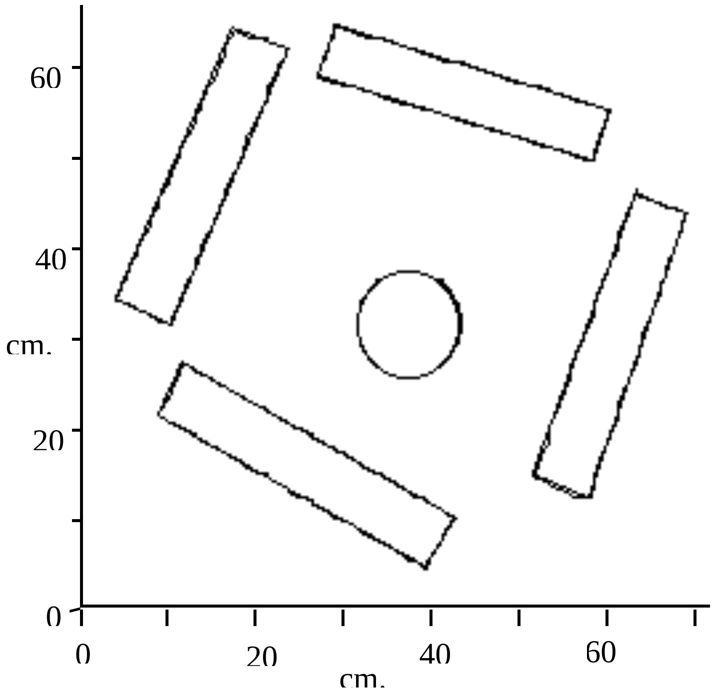

Figure 10b: An Aerial View of the Robot

Concluding Remarks

In the paper accompanying his 1991 Computers and Thought Award Lecture, Brooks remarked that,

> [The field of Knowledge Representation] concentrates much of its energies on anomalies within formal systems which are never used for any practical task.
>
> [Brooks, 1991a, page 578]

This paper should be construed as an answer to Brooks. According to the logical account given in this paper, a robot's incoming sensor data is filtered through an abductive process based on a framework of innate concepts, namely space, time, and causality.[20] The development of a rigorous, formal account of this process bridges the gap between theoretical work in Knowledge Representation and practical work in robotics, and opens up a great many possibilities for further research. The following issues, in ascending order of importance, are particularly pressing.

---

[20] This is somewhat reminiscent of Kant, according to whom, "the natural world as we know it, the whole content of our experience, is thoroughly conditioned by [certain] features: our experience is essentially experience of a spatio-temporal world of law-governed objects conceived of as distinct from our temporally successive experiences of them" [Strawson, 1966, page 21].

- The development of a more qualitative approach to spatial representation, possibly using a framework like that described by Randell *et al.* [1992], but augmented with the ability to handle certain kinds of metric information.
- The assimilation of sensor data from moving objects, such as humans, animals, or other robots. Movable obstacles should also be on the agenda.
- The assimilation of richer sensor data than that supplied by the Rug Warrior's simple bump switches.
- The control of the robot via the model of the world it acquires through abduction.

The last of these issues, namely robot control, deserves some comment. Although there's nothing particularly remarkable about the map-building capabilities of the Rug Warrior implementation described in Section 10, which serves mainly to illustrate the theoretical and methodological ideas whose promotion is the paper's main purpose, its design does suggest a sharply bipartite architecture which differs markedly from those hinted at in most current work in the Cognitive Robotics vein [Lespérance, *et al.*, 1994], [Kowalski, 1995], [Poole, 1995].

The architecture of the Rug Warrior implementation is behaviour-based in the purest sense, and yet serves the purpose of building a symbolic model of the world for which a rigorous denotational account can be supplied. In a sense, this architecture accommodates two extremes in robot design — the behaviour-based approach and the Cognitive Robotics approach — but makes no concessions to do so. At present, the world model constructed by the robot is a passive by-product of its activity. But the purity of this bipartite architecture can be maintained when the model plays a more active role.

This is accomplished by using the model to inform the decisions of a rational core which attempts to orchestrate the talents the robot has at its disposal, in the form of a repertoire of Brooks-style behaviours, in order to achieve its goals. The behaviour-based component of the architecture needn't be under the direct control of the rational component, but can remain fully autonomous. The aim of the rational, goal-achieving component of the architecture is only to perturb these behaviours in ways conducive to the achievement of its goals. Although the very idea of a goal to some extent compromises the behaviour-based approach to robotics, this proposal preserves one of its essential insights, which is that tight coupling between sensors and effectors is the way to achieve robustness in control.

Acknowledgements

The inspiration for Cognitive Robotics comes from Ray Reiter and his colleagues at the University of Toronto. Neelakantan Kartha patiently explained his formalisation of actions to me at IJCAI 95 over many cups of tea, and this led to the frame problem solution presented in Section 2. Thanks to Rob Miller for careful proof reading and valuable feedback. Thanks to Peter Cheung of the Department of Electrical Engineering at Imperial College for financial support for the purchase of the robot. Finally, thanks to Chitta Baral and Paulo Santos for their questions and suggestions during the ETAI discussion period. Most of the work reported in this paper was undertaken while the author was an EPSRC Advanced Research Fellow.

References

[Bacchus, *et al.*, 1995] F.Bacchus, J.Y.Halpren, and H.J.Levesque, Reasoning about Noisy Sensors in the Situation Calculus, *Proceedings IJCAI 95*, pp. 1933–1940.

[Baker, 1989], A.B.Baker, A Simple Solution to the Yale Shooting Problem, *Proceedings 1989 Knowledge Representation Conference*, pp. 11–20.

[Brooks, 1986] R.A.Brooks, A Robust Layered Control System for a Mobile Robot, *IEEE Journal of Robotics and Automation*, vol 2, no 1 (1986), pp. 14–23.

[Brooks, 1991a] R.A.Brooks, Intelligence Without Reason, *Proceedings IJCAI 91*, pp. 569–595.

[Brooks, 1991b] R.A.Brooks, Artificial Life and Real Robots, *Proceedings of the First European Conference on Artificial Life (1991)*, pp 3–10.

[Charniak & McDermott, 1985] E.Charniak and D.McDermott, *Introduction to Artificial Intelligence*, Addison-Wesley (1985).

[Davis, 1990] E.Davis, *Representations of Commonsense Knowledge*, Morgan Kaufmann (1990).

[Hanks & McDermott, 1987] S.Hanks and D.McDermott, Nonmonotonic Logic and Temporal Projection, *Artificial Intelligence*, vol 33 (1987), pp. 379–412.

[Harnad, 1990] S.Harnad, The Symbol Grounding Problem, *Physica D*, vol 42 (1990), pp. 335–346.

[Hayes, 1985] P.J.Hayes, The Second Naive Physics Manifesto, in *Formal Theories of the Commonsense World*, ed. Hobbs J.R. and Moore R.C., Ablex (1985), pp. 1–36.

[Jones & Flynn, 1993] J.L.Jones and A.M.Flynn, *Mobile Robots: Inspiration to Implementation*, A.K.Peters (1993).

[Kakas, Kowalski & Toni, 1993] A.C.Kakas, R.A.Kowalski and F.Toni, Abductive Logic Programming, *Journal of Logic and Computation*, vol 2, no 6 (1993), pp. 719–770.

[Kartha & Lifschitz, 1994] G.N.Kartha and V.Lifschitz, Actions with Indirect Effects (Preliminary Report), *Proceedings 1994 Knowledge Representation Conference (KR 94)*, pp. 341–350.

[Kartha & Lifschitz, 1995] G.N.Kartha and V.Lifschitz, A Simple Formalization of Actions Using Circumscription, *Proceedings IJCAI 95*, pp. 1970–1975.

[Kowalski, 1995] R.A.Kowalski, Using Meta-Logic to Reconcile Reactive with Rational Agents, in *Meta-Logics and Logic Programming*, ed. K.R.Apt and F.Turini, MIT Press (1995), pp. 227–242.

[Kowalski & Sergot, 1986] R.A.Kowalski and M.J.Sergot, A Logic-Based Calculus of Events, *New Generation Computing*, vol 4 (1986), pp. 67–95.

[Lespérance, *et al.*, 1994] Y.Lespérance, H.J.Levesque, F.Lin, D.Marcu, R.Reiter, and R.B.Scherl, A Logical Approach to High-Level Robot Programming: A Progress Report, in *Control of the Physical World by Intelligent Systems: Papers from the 1994 AAAI Fall Symposium*, ed. B.Kuipers, New Orleans (1994), pp. 79–85.

[Lifschitz, 1994] V.Lifschitz, Circumscription, in *The Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*, ed. D.M.Gabbay, C.J.Hogger and J.A.Robinson, Oxford University Press (1994), pp. 297–352.

[McCarthy, 1986] J.McCarthy, Applications of Circumscription to Formalizing Common Sense Knowledge, *Artificial Intelligence*, vol 26 (1986), pp. 89–116.

[McCarthy, 1989] J.McCarthy, Artificial Intelligence, Logic and Formalizing Common Sense, in *Philosophical Logic and Artificial Intelligence*, ed. R.Thomason, Kluwer Academic (1989), pp. 161–190.

[Miller, 1996] R.S.Miller, A Case Study in Reasoning about Actions and Continuous Change, *Proceedings ECAI 96*, pp. 624–628.

[Nilsson, 1984] N.J.Nilsson, ed., *Shakey the Robot*, SRI Technical Note no. 323 (1984), SRI, Menlo Park, California.

[Poole, 1995] D.Poole, Logic Programming for Robot Control, *Proceedings IJCAI 95*, pp. 150–157.

[Randell, *et al.*, 1992] D.A.Randell, Z.Cui and A.G.Cohn, A Spatial Logic Based on Regions and Connection, *Proceedings 1992 Knowledge Representation Conference (KR 92)*, pp. 165–176.

[Reiter, 1987] R.Reiter, A Theory of Diagnosis from First Principles, *Artificial Intelligence*, vol 32 (1987), pp. 57–95.

[Reiter, 1996] R.Reiter, Natural Actions, Concurrency and Continuous Time in the Situation Calculus, *Proceedings 1998 Knowledge Representation Conference (KR 98)*, pp. 176–185.

[Sandewall, 1989] E.Sandewall, Combining Logic and Differential Equations for Describing Real-World Systems, *Proceedings 1989 Knowledge Representation Conference (KR 89)*, pp. 412–420.

[Sandewall, 1994] E.Sandewall, *Features and Fluents: The Representation of Knowledge about Dynamical Systems, Volume 1*, Oxford University Press, 1994.

[Shanahan, 1989] M.P.Shanahan, Prediction Is Deduction but Explanation Is Abduction, *Proceedings IJCAI 89*, pp. 1055–1060.

[Shanahan, 1990] M.P.Shanahan, Representing Continuous Change in the Event Calculus, *Proceedings ECAI 90*, pp. 598–603.

[Shanahan, 1995a] M.P.Shanahan, Default Reasoning about Spatial Occupancy, *Artificial Intelligence*, vol 74 (1995), pp. 147–163.

[Shanahan, 1995b] M.P.Shanahan, A Circumscriptive Calculus of Events, *Artificial Intelligence*, vol 77 (1995), pp. 249–284.

[Shanahan, 1996a] M.P.Shanahan, Robotics and the Common Sense Informatic Situation, *Proceedings ECAI 96*, pp. 684–688.

[Shanahan, 1996b] M.P.Shanahan, Noise and the Common Sense Informatic Situation for a Mobile Robot, *Proceedings AAAI 96*, pp. 1098–1103.

[Shanahan, 1997a] M.P.Shanahan, *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*, MIT Press (1997).

[Shanahan, 1997b] M.P.Shanahan, Noise, Non-Determinism and Spatial Uncertainty, *Proceedings AAAI 97*, pp. 153–158.

[Shanahan, 1998] M.P.Shanahan, Reinventing Shakey, *Working Notes of the 1998 AAAI Fall Symposium on Cognitive Robotics*, pp. 125–135.

[Strawson, 1966] P.F.Strawson, *The Bounds of Sense*, Methuen (1966).

Appendix A: Proofs of Propositions

Proof of Proposition 2.8. From CIRC[N ; Happens], we get,

$$\text{Happens}(a,t) \leftrightarrow [[a = \text{Rotate}(90) \wedge t = 10] \vee [a = \text{Rotate}(-180) \wedge t = 20]]. \quad \text{[A.1]}$$

From CIRC[E ; Initiates, Terminates, Releases], we get,

$$\begin{aligned}&\text{Initiates}(a,f,t) \leftrightarrow && \text{[A.2]}\\ &\quad \exists\, r1,r2,r3\, [a = \text{Rotate}(r1) \wedge f = \text{Facing}(r2) \wedge \\ &\qquad \text{HoldsAt}(\text{Facing}(r3),t) \wedge r2 = r3 + r1]\end{aligned}$$

$$\begin{aligned}&\text{Terminates}(a,f,t) \leftrightarrow && \text{[A.3]}\\ &\quad \exists\, r1,r2\, [a = \text{Rotate}(r1) \wedge f = \text{Facing}(r2) \wedge \\ &\qquad \text{HoldsAt}(\text{Facing}(r2),t) \wedge r1 \neq 0]\end{aligned}$$

$$\neg \exists\, a,f,t\, [\text{Releases}(a,f,t)]. \quad \text{[A.4]}$$

From [A.1] and (EC5), we get,

$$\neg\, \text{Clipped}(0,\text{Facing}(0),t) \leftarrow 0 \leq t \leq 10$$

which, given (EC1) and (2.3), yields,

$\text{HoldsAt(Facing(0),t)} \leftarrow 0 \leq t \leq 10.$ [A.5]

From [A.2] and [A.5], we get,

$\text{Initiates(Rotate(90),Facing(90),10).}$ [A.6]

From [A.1] and (EC5), we get,

$\neg\, \text{Clipped(10,Facing(90),t)} \leftarrow 10 < t \leq 20$

which, given [A.1], [A.6] and (EC2), yields,

$\text{HoldsAt(Facing(90),t)} \leftarrow 0 \leq t \leq 10.$ [A.7]

From [A.2] and [A.5], we get,

$\text{Initiates(Rotate(–180),Facing(270),20).}$ [A.8]

From [A.1] and (EC5), we get,

$\neg\, \text{Clipped(20,Facing(270),t)} \leftarrow 20 < t$

which, given [A.1], [A.8] and (EC2), yields,

$\text{HoldsAt(Facing(270),t)} \leftarrow 20 < t.$ [A.9]

The proposition follows from [A.5], [A.7], and [A.9]. □

Proof of Proposition 3.13. From CIRC[N ; Happens], we get,

$\text{Happens(a,t)} \leftrightarrow [[a = \text{Go} \wedge t = 10] \vee [a = \text{Stop} \wedge t = 20]].$ [A.10]

From CIRC[E ; Initiates, Terminates, Releases], we get,

$$\begin{aligned}&\text{Initiates(a,f,t)} \leftrightarrow \\ &\quad [a = \text{Go} \wedge f = \text{Moving}] \vee \\ &\quad\quad \exists\, p\, [a = \text{Stop} \wedge f = \text{Location(Robot,p)} \wedge \\ &\quad\quad\quad \text{HoldsAt(Location(Robot,p),t)}]]\end{aligned}$$ [A.11]

$\text{Terminates(a,f,t)} \leftrightarrow a = \text{Stop} \wedge f = \text{Moving}$ [A.12]

$\text{Releases(a,f,t)} \leftrightarrow \exists\, p\, [a = \text{Go} \wedge f = \text{Location(Robot,p)}].$ [A.13]

From [A.10] and (EC5), we get,

$\neg\, \text{Clipped}(0,\text{Location(Robot},\langle 0,0\rangle),t) \leftarrow 0 \leq t \leq 10$

which, from (EC1) and (3.9), yields,

$\text{HoldsAt(Location(Robot},\langle 0,0\rangle),t) \leftarrow 0 \leq t \leq 10.$ [A.14]

Similarly, we can show,

$\text{HoldsAt(Facing(90),10).}$ [A.15]

From [A.11] we have,

$\text{Initiates(Go,Moving,10).}$ [A.16]

From [A.10] and (EC5), we get,

$\neg$ Clipped(10,Moving,t) $\leftarrow$ 10 < t $\leq$ 20. [A.17]

From [A.14], [A.15], and (3.5), we get,

Trajectory(Moving,10,Location(Robot,$\langle$x,0$\rangle$),d) $\leftarrow$ x = V.d

which, given [A.10], [A.16], [A.17] and (EC7), yields,

HoldsAt(Location(Robot,$\langle$x,0$\rangle$),t) $\leftarrow$ 10 < t $\leq$ 20 $\wedge$ x = V.(t – 10). [A.18]

From [A.11] and [A.18], we have,

Initiates(Stop,Location(Robot,$\langle$x,0$\rangle$),20) $\leftarrow$ x = V.10. [A.19]

From [A.10] and (EC5), we get,

$\neg$ Clipped(20,Location(Robot,$\langle$x,0$\rangle$),t) $\leftarrow$ 20 < t $\wedge$ x = V.10

which, given [A.10], [A.18], [A.19] and (EC2), yields,

HoldsAt(Location(Robot,$\langle$x,0$\rangle$),t) $\leftarrow$ 20 < t $\wedge$ x = V.10. [A.20]

From [A.14], [A.18] and [A.20], we arrive at,

HoldsAt(Location(Robot,$\langle$x,y$\rangle$),t) $\leftarrow$
[0 $\leq$ t $\leq$ 10 $\wedge$ x = 0 $\wedge$ y = 0] $\vee$
[10 < t $\leq$ 20 $\wedge$ x = V.(t – 10) $\wedge$ y = 0] $\vee$
[20 < t $\wedge$ x = V.10 $\wedge$ y = 0].

The proposition follows from this and the domain constraint (3.6). □

Proof of Proposition 5.8. From CIRC[N1 $\wedge$ N2 ; Happens], we get,

Happens(a,t) $\leftrightarrow$ H1(a,t) $\vee$ H2(a,t) $\vee$ H3(a,t) $\vee$ H4(a,t) [A.21]

where,

H1(a,t) $\equiv_{def}$
[a = Go $\wedge$ t = 0] $\vee$ [a = Stop $\wedge$ t = 2·8] $\vee$
[a = Rotate(–90) $\wedge$ t = 3·3] $\vee$ [a = Go $\wedge$ t = 3·8]

H2(a,t) $\equiv_{def}$
$\exists$ w,r [a = Bump $\wedge$ [HoldsAt(Moving,t) $\vee$ t = 0 $\vee$ t = 3·8] $\wedge$
HoldsAt(Facing(r),t) $\wedge$ HoldsAt(Blocked(Robot,w,r),t)]

H3(a,t) $\equiv_{def}$
$\exists$ w,r,p1,p2,z [a = Switch1 $\wedge$ [HoldsAt(Moving,t) $\vee$ t = 0 $\vee$ t = 3·8] $\wedge$
HoldsAt(Facing(r),t) $\wedge$
HoldsAt(Occupies(Robot,Displace(Disc(z),p1)),t) $\wedge$
HoldsAt(Touching(Robot,w,p2),t) $\wedge$
r–150 < Bearing(p1,p2) < r+30]

$H4(a,t) \equiv_{def}$

$$\exists\, w,r,p1,p2,z\ [a = Switch2 \wedge [HoldsAt(Moving,t) \vee t = 0 \vee t = 3{\cdot}8] \wedge$$
$$HoldsAt(Facing(r),t) \wedge$$
$$HoldsAt(Occupies(Robot,Displace(Disc(z),p1)),t) \wedge$$
$$HoldsAt(Touching(Robot,w,p2),t) \wedge$$
$$r{-}30 < Bearing(p1,p2) < r{+}150].$$

From CIRC[E ; Initiates, Terminates, Releases], we get,

$$Initiates(a,f,t) \leftrightarrow \qquad [A.22]$$
$$\exists\, r1,r2\ [a = Rotate(r1) \wedge f = Facing(r1+r2) \wedge HoldsAt(Facing(r2),t)] \vee$$
$$[a = Go \wedge f = Moving] \vee$$
$$\exists\, g\ [[a = Stop \vee \exists\, r\ [a = Bump(r)]] \wedge f = Occupies(Robot,g) \wedge$$
$$HoldsAt(Occupies(Robot,g),t)]$$

$$Terminates(a,f,t) \leftrightarrow \qquad [A.23]$$
$$[a = Stop \vee \exists\, r\ [a = Bump(r) \vee a = Rotate(r)]] \wedge f = Moving$$

$$Releases(a,f,t) \leftrightarrow \exists\, g\ [a = Go \wedge f = Occupies(Robot,g)]. \qquad [A.24]$$

From CIRC [$O \wedge M1 \wedge M2$; AbSpace ; $Initially_P$, $Initially_N$] we get,

$$Initially_N(Occupies(x,g)) \leftarrow x \neq A \wedge x \neq Robot. \qquad [A.25]$$

It can easily be shown that A retains its initial region of occupancy for all time. Let $X_{turn} = 1 + 2{\cdot}8.Sin(80)$, $Y_{turn} = 1 + 2{\cdot}8.Cos(80)$. From [A.25] and (5.7), using Axioms (Sp7) and (B5), it can be confirmed that,

$$\neg\, HoldsAt(Blocked(Robot,w,r),t) \leftarrow \qquad [A.26]$$
$$HoldsAt(Occupies(Robot,Displace(Disc(0{\cdot}5),\langle x,y\rangle)),t) \wedge$$
$$\exists\, d\ [[0 \leq d \leq 2{\cdot}8 \wedge x = 1 + d.Sin(80) \wedge y = 1 + d.Cos(80)] \vee$$
$$[3{\cdot}8 < d < T_{bump} \wedge x = X_{turn} + (d - 3{\cdot}8).Sin(-10) \wedge$$
$$y = Y_{turn} + (d - 3{\cdot}8).Cos(-10)]].$$

It can similarly be confirmed that,

$$\neg\, HoldsAt(Touching(Robot,w,p),t) \leftarrow \qquad [A.27]$$
$$HoldsAt(Occupies(Robot,Displace(Disc(0{\cdot}5),\langle x,y\rangle)),t) \wedge$$
$$\exists\, d\ [[0 \leq d \leq 2{\cdot}8 \wedge x = 1 + d.Sin(80) \wedge y = 1 + d.Cos(80)] \vee$$
$$[3{\cdot}8 < d < T_{bump} \wedge x = X_{turn} + (d - 3{\cdot}8).Sin(-10) \wedge$$
$$y = Y_{turn} + (d - 3{\cdot}8).Cos(-10)]].$$

Given [A.26], from [A.21] to [A.24], using a similar procedure to that employed in the proof of Proposition 3.13, we can show,

$$\text{HoldsAt(Occupies(Robot,Displace(Disc}(0{\cdot}5),\langle x,y\rangle)),t) \leftarrow \quad [A.28]$$
$$[0 \le t \le 2{\cdot}8 \wedge x = 1 + t.\mathrm{Sin}(80) \wedge y = 1 + t.\mathrm{Cos}(80)] \vee$$
$$[2{\cdot}8 < t \le 3.8 \wedge x = X_{turn} \wedge y = Y_{turn}] \vee$$
$$[3{\cdot}8 < t \le T_{bump} \wedge x = X_{turn} + (t - 3{\cdot}8).\mathrm{Sin}(-10) \wedge$$
$$y = Y_{turn} + (t - 3{\cdot}8).\mathrm{Cos}(-10)].$$

Given that A retains its initial region of occupancy, from [A.28], [A.25] and (5.7), using Axioms (Sp7), we can show,

$$\text{HoldsAt(Blocked(Robot,A,}-10),T_{bump}). \quad [A.29]$$

We can also show,

$$\text{HoldsAt(Facing(}-10),T_{bump}). \quad [A.30]$$

From [A.29] and [A.30], using Axiom (B5), we get,

$$\text{Happens(Bump,}T_{bump}). \quad [A.31]$$

Given that A retains its initial region of occupancy, from [A.28], [A.25] and (5.7), using Axioms (Sp7), we can also show,

$$\exists\ p1, p2\ [\text{HoldsAt(Touching(Robot,A,p1),}T_{bump}) \wedge \quad [A.32]$$
$$\text{HoldsAt(Occupies(Robot,Displace(Disc}(0{\cdot}5),p2)),T_{bump}) \wedge$$
$$\text{Bearing}(p1,p2) = 0].$$

From [A.21] and [A.30] to [A.32] we get,

$$\text{Happens(Switch1,}T_{bump}) \wedge \text{Happens(Switch2,}T_{bump}). \quad [A.33]$$

From [A.21], [A.27] and Axiom (B6) we get,

$$[\text{Happens(Switch1,t)} \vee \text{Happens(Switch2,t)}] \rightarrow t = T_{bump}. \quad [A.34]$$

The proposition follows directly from [A.33] and [A.34]. □

Proof of Proposition 6.2. The proof is analogous to that of Proposition 5.10. □

Proof of Proposition 7.9. The proposition can be proved using Theorem 8.9. First, it can easily be verified that O $\wedge$ M1 $\wedge$ M2 is consistent. Now, let ρ_0 be $\langle 1,1\rangle$, let ρ_1 be $\langle 2,3\rangle$, and let τ_1 be $2{\cdot}1$. We have $T_{stop}(1) = \tau_1$, $RL(\rho_0,0,1) = \langle 2,3{\cdot}1\rangle$, and $\varepsilon.MT(0,1) = 0{\cdot}525$. Therefore the distance from ρ_1 to $RL(\rho_0,0,1)$ is less than $\varepsilon.MT(0,1)$, and the sequence comprising just ρ_1 is a bump trail.

Let η_1 be 3. We have $RO(\tau_1) = 0$, which gives us the following.

$$O \wedge M1 \wedge M2 \vDash OTR(\rho_1,BD_{min}(\eta_1,RO(\tau_1)),BD_{max}(\eta_1,RO(\tau_1)))$$

This follows from the definition of OTR, since the obstacle A would touch the robot at point $\langle 2,3{\cdot}5\rangle$ if the robot's location were $\rho_1 = \langle 2,3\rangle$.

Finally, it can be shown that,

$\text{CIRC}[\text{O} \wedge \text{M1} \wedge \text{M2} \text{ ; AbSpace ; Initially}_\text{P}, \text{Initially}_\text{N}] \wedge$
$\quad \text{CIRC}[\text{N1} \wedge \text{N2} \text{ ; Happens}] \wedge$
$\quad\quad \text{CIRC}[\text{E ; Initiates, Terminates, Releases}] \wedge \text{B} \nvDash \neg\, \text{COMP}[\Psi]$

because a model can be exhibited which satisfies,

$[\text{Happens}(a,t) \wedge [a = \text{Switch1} \vee a = \text{Switch2}]] \rightarrow t = 2{\cdot}1.$

The proposition then follows from Theorem 8.9. □

Appendix B: Proofs of Theorems

Proof of Theorem 7.13. We only need to consider Ψ since the definition of an explanation caters for COMP[Ψ] automatically. The theorem follows from the fact that Axioms (B3) and (B6) are equivalent if ε is 0, and the fact that (B3) ensures that the robot's path is deterministic in the sense that at any given time its location is the same in every model of,

$\text{CIRC}[\text{O} \wedge \text{M1} \wedge \text{M2} \text{ ; AbSpace ; Initially}_\text{P}, \text{Initially}_\text{N}] \wedge$
$\quad \text{CIRC}[\text{N1} \wedge \text{N2} \text{ ; Happens}] \wedge$
$\quad\quad \text{CIRC}[\text{E ; Initiates, Terminates, Releases}] \wedge \text{B}_\text{det}.$

To see that the theorem follows, consider that, if the robot's path is deterministic according to a formula Γ and the locations and shapes of objects are the same in every model of Γ (as they must be in the above formula since M1 and M2 are complete spatial descriptions), then $\Gamma \nvDash \neg\, \Psi$ if and only if $\Gamma \vDash \Psi$. □

For the proofs of Theorems 8.8 to 8.10, let Σ be,

$\text{CIRC}[\text{O} \wedge \text{M1} \wedge \text{M2} \text{ ; AbSpace ; Initially}_\text{P}, \text{Initially}_\text{N}] \wedge$
$\quad \text{CIRC}[\text{N1} \wedge \text{N2} \text{ ; Happens}] \wedge$
$\quad\quad \text{CIRC}[\text{E ; Initiates, Terminates, Releases}] \wedge \text{B}.$

The proof of Theorem 8.8 requires the development of two lemmas.

First we note that from CIRC[N1 $\wedge$ N2 ; Happens], we get the completion of Happens (as in the proof of Proposition 5.9, for example). If N2 tracks Ψ, then this completion will include a Stop action for every time point at which Moving is terminated. So the robot is sure to remain in continuous motion between temporally adjacent Go and Stop actions.

Lemma B.1. For any model V of Σ,

$\text{V} \Vdash \text{HoldsAt}(\text{Facing}(\theta),\tau)$

if and only if $\theta = \text{RO}(\tau)$.

Proof. Given the completion of Happens, the proof is a straightforward induction on the number of Rotate actions in N2, using Axioms (E1) and (E2), and the event calculus axioms (EC1) to (EC7). □

Lemma B.2. If N2 tracks Ψ, then for any model V of Σ and any i, i < 0 < m, if

$$V \Vdash \text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),\rho)),\tau_i)]$$

then,

$$V \Vdash \exists\, p\ [\text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),p)),\tau_{i+1}) \wedge \text{Distance}(p,\text{RL}(\rho,j,k)) \leq \varepsilon.\text{MT}(j,k)]$$

where $T_{stop}(j)$ is τ_i and $T_{stop}(k)$ is τ_{i+1}.

Proof. What we in fact show is that, for any i, $0 < i \leq n - j$, if,

$$V \Vdash \text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),\rho)),T_{stop}(j))] \qquad \text{[B.3]}$$

then,

$$V \Vdash \exists\, p\ [\text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),p)),T_{stop}(j+i)) \wedge \text{Distance}(p,\text{RL}(\rho,j,j+i)) \leq \varepsilon.\text{MT}(j,j+i)]$$

from which the lemma follows directly, given that N2 tracks Ψ, since $k = j + i$ for some i, $0 < i \leq n - j$.

The proof is by induction over i. The base case is as follows. There is exactly one Go action between $T_{stop}(j)$ and $T_{stop}(j+1)$, which occurs at time $T_{go}(j+1)$. From the completion of Happens, the event calculus axioms, and [B.3], we can show that,

$$V \Vdash \text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),\rho)),T_{go}(j+1))].$$

Now given that,

$$V \Vdash \text{HoldsAt}(\text{Facing}(\theta),T_{go}(j+1))]$$

(which, from Lemma B.1, is equivalent to $\theta = \text{RO}(T_{go}(j+1)))$, we can show,

$$V \Vdash \exists\, p\ [\text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),p)),T_{stop}(j+1)) \wedge \text{Distance}(p,\rho') \leq \varepsilon.(T_{stop}(j+1) - T_{go}(j+1))]$$

where ρ' is,

$$\rho + \langle (T_{stop}(j+1) - T_{go}(j+1)).\text{Sin}(\theta),(T_{stop}(j+1) - T_{go}(j+1)).\text{Cos}(\theta)\rangle.$$

Given the completion of Happens, this follows from Axioms (E3) and (B6), and the event calculus axioms (EC1) to (EC7).

From the definitions of RL and MT, this implies,

$$V \Vdash \exists\, p\ [\text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),p)),T_{stop}(j+1)) \wedge \text{Distance}(p,\text{RL}(\rho,j,j+1)) \leq \varepsilon.\text{MT}(j,j+1)].$$

The inductive step is as follows. Consider any h, $0 < h \leq n - j$. Suppose we know that if,

$$V \Vdash \text{HoldsAt}(\text{Occupies}(\text{Robot},\text{Displace}(\text{Disc}(0{\cdot}5),\rho)),T_{stop}(j))] \qquad \text{[B.4]}$$

then,

$$V \Vdash \exists p\ [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),T_{stop}(j+h)) \wedge \\ \mathrm{Distance}(p,\mathrm{RL}(\rho,j,j+h)) \leq \varepsilon.\mathrm{MT}(j,j+h)].$$

Then what's required is to show that if [B.4] then,

$$V \Vdash \exists p\ [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),T_{stop}(j+h+1)) \wedge \\ \mathrm{Distance}(p,\mathrm{RL}(\rho,j,j+h+1)) \leq \varepsilon.\mathrm{MT}(j,j+h+1)].$$

The proof is similar to that for the base case. From the completion of Happens, the event calculus axioms, and [B.4], we can show that,

$$V \Vdash \exists p\ [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),T_{go}(j+h+1)) \wedge \\ \mathrm{Distance}(p,\mathrm{RL}(\rho,j,j+h)) \leq \varepsilon.\mathrm{MT}(j,j+h)].$$

Now given that,

$$V \Vdash \mathrm{HoldsAt}(\mathrm{Facing}(\theta),T_{go}(j+h+1))]$$

(which, from Lemma B.1, is equivalent to $\theta = \mathrm{RO}(T_{go}(j+h+1)))$, we can show,

$$V \Vdash \exists p\ [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),T_{stop}(j+h+1)) \wedge \\ \mathrm{Distance}(p,\rho') \leq \varepsilon.(T_{stop}(j+h+1) - T_{go}(j+h+1))]$$

where ρ' is,

$$\rho + \langle (T_{stop}(j+h+1) - T_{go}(j+h+1)).\mathrm{Sin}(\theta),(T_{stop}(j+h+1) - T_{go}(j+h+1)).\mathrm{Cos}(\theta)\rangle.$$

Given the completion of Happens, this follows from Axioms (E3) and (B6), and the event calculus axioms (EC1) to (EC7).

From the definitions of RL and MT, this implies,

$$V \Vdash \exists p\ [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),T_{stop}(j+h+1)) \wedge \\ \mathrm{Distance}(p,\mathrm{RL}(\rho,j,j+h+1)) \leq \varepsilon.\mathrm{MT}(j,j+h+1)]. \qquad \square$$

Proof of Theorem 8.8. The proof is by induction on the elements of the bump trail. The base case is as follows. Let $T_{stop}(k)$ be τ_1 and ρ_0 be the robot's initial location according to M1. Consider ψ_1. Suppose that there is no ρ_1 such that,

- the distance from ρ_1 to the point $\mathrm{RL}(\rho_0,0,k)$ is less than or equal to $\varepsilon.\mathrm{MT}(0,k)$, and
- $O \wedge M1 \wedge M2 \vDash \mathrm{OTR}(\rho_1,\mathrm{BD}_{min}(\eta_1,\mathrm{RO}(\tau_1)),\mathrm{BD}_{max}(\eta_1,\mathrm{RO}(\tau_1)))$.

So, for any ρ_1 whose distance from $\mathrm{RL}(\rho_0,0,k)$ is less than or equal to $\varepsilon.\mathrm{MT}(0,k)$, we have,

$$O \wedge M1 \wedge M2 \nvDash \mathrm{OTR}(\rho_1,\mathrm{BD}_{min}(\eta_1,\mathrm{RO}(\tau_1)),\mathrm{BD}_{max}(\eta_1,\mathrm{RO}(\tau_1))).$$

Clearly we have,

$$\mathrm{CIRC}[O \wedge M1 \wedge M2\ ;\ \mathrm{AbSpace}\ ;\ \mathrm{Initially}_P,\ \mathrm{Initially}_N] \nvDash \\ \mathrm{OTR}(\rho_1,\mathrm{BD}_{min}(\eta_1,\mathrm{RO}(\tau_1)),\mathrm{BD}_{max}(\eta_1,\mathrm{RO}(\tau_1)))$$

since the circumscription only cuts out models with superflous objects. But we also have,

$$\Sigma \not\models \mathrm{OTR}(\rho_1,\mathrm{BD}_{min}(\eta_1,\mathrm{RO}(\tau_1)),\mathrm{BD}_{max}(\eta_1,\mathrm{RO}(\tau_1))).$$

This follows from the fact that the initial value of the spatial fluent Occupies is fixed by the first circumscription in Σ and are not influenced by the rest of Σ.

So there is some model V of Σ such that,

$$V \Vdash \neg\, \mathrm{OTR}(\rho_1,\mathrm{BD}_{min}(\eta_1,\mathrm{RO}(\tau_1)),\mathrm{BD}_{max}(\eta_1,\mathrm{RO}(\tau_1))). \quad [B.5]$$

From Lemma B.2, V must satisfy,

$$\exists\, p\, [\mathrm{HoldsAt}(\mathrm{Occupies}(\mathrm{Robot},\mathrm{Displace}(\mathrm{Disc}(0{\cdot}5),p)),\tau_1) \wedge \mathrm{Distance}(p,\mathrm{RL}(\rho_0,0,k)) \leq \varepsilon.\mathrm{MT}(0,k)]. \quad [B.6]$$

From [B.5] and [B.6] and the completion of Happens, given that the distance from ρ_1 to the point $\mathrm{RL}(\rho_0,0,k)$ is less than or equal to $\varepsilon.\mathrm{MT}(0,k)$, we can show,

$$V \Vdash \neg\, \mathrm{Happens}(\mathrm{Switch1},\tau_1)$$

if η_1 is 1 or 3, and,

$$V \Vdash \neg\, \mathrm{Happens}(\mathrm{Switch2},\tau_1)$$

if η_1 is 2 or 3. Therefore M2 is not an explanation of Ψ, which is a contradiction.

The inductive step is similar to the base case, and is as follows. Consider any h, $0 < h < m$. Let $T_{stop}(k)$ be τ_{h+1}. Suppose that there exists a bump trail ρ_1 to ρ_h such that, for any i, $0 < i \leq h$,

$$O \wedge M1 \wedge M2 \models \mathrm{OTR}(\rho_i,\mathrm{BD}_{min}(\eta_i,\mathrm{RO}(\tau_i)),\mathrm{BD}_{max}(\eta_i,\mathrm{RO}(\tau_i)))$$

for which there is no ρ_{h+1} such that,

- the distance from ρ_{h+1} to the point $\mathrm{RL}(\rho_h,h,k)$ is less than or equal to $\varepsilon.\mathrm{MT}(h,k)$, and
- $O \wedge M1 \wedge M2 \models \mathrm{OTR}(\rho_{h+1},\theta_{min},\theta_{max})$

where $\theta_{min} = \mathrm{BD}_{min}(\eta_{h+1},\mathrm{RO}(\tau_{h+1}))$ and $\theta_{max} = \mathrm{BD}_{max}(\eta_{h+1},\mathrm{RO}(\tau_{h+1}))$.

So, for any ρ_{h+1} whose distance from $\mathrm{RL}(\rho_h,h,k)$ is less than or equal to $\varepsilon.\mathrm{MT}(h,k)$, we have,

$$O \wedge M1 \wedge M2 \not\models \mathrm{OTR}(\rho_{h+1},\theta_{min},\theta_{max}).$$

Using the same argument as for the base case, this means there is some model V of Σ such that,

$$V \Vdash \neg\, \mathrm{OTR}(\rho_{h+1},\theta_{min},\theta_{max}). \quad [B.7]$$

From Lemma B.2, V must satisfy,

$\exists$ p [HoldsAt(Occupies(Robot,Displace(Disc(0·5),p)),τ_{h+1}) $\wedge$
Distance(p,RL(ρ_h,h,k)) $\leq$ ε.MT(h,k)]. [B.8]

From [B.7] and [B.8] and the completion of Happens, given that the distance from ρ_{h+1} to the point RL(ρ_h,h,k) is less than or equal to ε.MT(h,k), we can show,

V $\Vdash \neg$ Happens(Switch1,τ_{h+1})

if η_{h+1} is 1 or 3, and,

V $\Vdash \neg$ Happens(Switch2,τ_{h+1})

if η_{h+1} is 2 or 3. Therefore M2 is not an explanation of Ψ, which is a contradiction. □

Proof of Theorem 8.9. Suppose,

$\Sigma \nvDash \neg$ COMP[Ψ] [B.9]

and let ρ_1 to ρ_m be a bump trail such that, for any i, 0 < i ≤ m,

O $\wedge$ M1 $\wedge$ M2 $\vDash$ OTR(ρ_i,BD_{min}(η_i,RO(τ_i)),BD_{max}(η_i,RO(τ_i))). [B.10]

Now, given [B.9], it's sufficient to exhibit a model of Σ such that V $\Vdash \psi_i$ for any ψ_i. The following is a partial definition of such a model V. The definition doesn't encompass predicate and function symbols whose interpretations are straightforward to fill in.

Let each ground fluent or action term be interpreted by itself. Let Happens be interpreted as follows.

V⟦Happens⟧ = V1 ∪ V2 ∪ V3

where V1 is,

{⟨α,τ⟩ | Happens(α,τ) occurs in N2},

V2 is,

{⟨α,τ⟩ | Happens(α,τ) occurs in some ψ_i},

V3 is,

{⟨Bump,τ⟩ | Happens(Switch1,τ) or Happens(Switch2,τ) occurs in some ψ_i}.

Obviously, from V1, V satisfies N2, and from V2, V satisfies Ψ. But we need to fill in the interpretation of HoldsAt for the fluents Facing and Blocked in such a way that V satisfies Axioms (H1) to (H3) given V1 to V3. The only-if counterparts to these axiom must also be satisfied if V is to be minimal with respect to Happens. To begin with, we incorporate in V the objects described by M1 and M2. To do this, we first supply interpretations for the function Shape and the predicate Initially_P.

Let V_O be any model of,

CIRC[O $\wedge$ M1 $\wedge$ M2 ; AbSpace ; Initially_P, Initially_N].

We know that $O \wedge M1 \wedge M2$ is consistent, so it must have at least one model. Given this, it's easy to see that it must have a minimal model with respect to the above circumscription policy. Let V_O be any such minimal model, and let Shape, Initially_P, and Initially_N be interpreted as follows.

$$V[\![\text{Shape}]\!] = V_O[\![\text{Shape}]\!]$$

$$V[\![\text{Initially}_P]\!] = \{\langle\beta\rangle \mid \langle\beta\rangle \in V_O[\![\text{Initially}_P]\!] \text{ and } \beta \text{ is Occupies}(\omega,\gamma) \text{ or Facing}(\theta)\}$$

$$V[\![\text{Initially}_N]\!] = \{\langle\beta\rangle \mid \langle\beta\rangle \in V_O[\![\text{Initially}_N]\!] \text{ and } \beta \text{ is Occupies}(\omega,\gamma) \text{ or Facing}(\theta)\}$$

$$V[\![\text{AbSpace}]\!] = V_O[\![\text{AbSpace}]\!]$$

Obviously V satisfies $O \wedge M1 \wedge M2$ and its circumscription.

Now we can construct a trajectory for the robot which brings it into contact with the obstacles described in M1 and M2 in such a way as to cause the sensor events described in Ψ.

For every i, $0 < i \leq n$ such that $T_{stop}(i) = \tau_j$ for some j, $0 < j \leq m$, let ρ_i* be ρ_j. Let every other ρ_i* be some point such that the distance from ρ_{i+1}* to $RL(\rho_i*,i,i+1)$ is less than or equal to $\varepsilon.MT(i,i+1)$. From the definition of a bump trail, such an assignment of points must exist.

The point ρ_i* is the location in V of the robot at the time of the i^{th} Stop action. Some Stop actions will be associated with Bump events and some won't. For those that are, the choice of ρ_i* ensures that the robot abuts an object, since from [B.10] we know that, for any $0 < j \leq m$,

$$V \Vdash OTR(\rho_j, BD_{min}(\eta_j, RO(\tau_j)), BD_{max}(\eta_j, RO(\tau_j))).$$

Now, HoldsAt is interpreted as follows.

$$V[\![\text{HoldsAt}]\!] = V4 \cup V5 \cup V6 \cup V7$$

where V4 to V7 are defined in the following way. V4 and V5 interpret the Occupies fluent. V4 describes the regions of occupancy of all objects except the robot, which is the only object which moves.

$$V4 = \{\langle\text{Occupies}(\omega,\gamma),\tau\rangle \mid \omega \neq \text{Robot and } V \Vdash \text{Initially}_P(\text{Occupies}(\omega,\gamma))\}$$

V5 describes the robot's region of occupancy. The idea here is to plot a course between each ρ_i* and ρ_{i+1}*. Let θ_i and δ_i satisfy the following equation.

$$\rho_{i+1}* = \rho_i* + \langle\delta_i.\text{Sin}(\theta_i),\delta_i.\text{Cos}(\theta_i)\rangle$$

The LOC function, defined below, captures the robot's location between Stop actions.

$$LOC(\tau) = \begin{cases} \rho \text{ if } \tau \leq T_{go}(1), \\ LOC(T_{go}(i)) + DISP(i,\tau) \text{ if } T_{go}(i) < \tau \leq T_{stop}(i), \\ LOC(T_{stop}(i)) \text{ if } T_{go}(i+1) \text{ is defined and } T_{stop}(i) < \tau \leq T_{go}(i+1), \\ LOC(T_{stop}(i)) \text{ if } T_{go}(i+1) \text{ is undefined and } \tau > T_{stop}(i) \end{cases}$$

where

- $V \Vdash \text{Initially}_P(\text{Occupies}(\text{Robot},\gamma))$ and $\gamma = \text{Displace}(\text{Disc}(0{\cdot}5),\rho))$, and,
- $\text{DISP}(i,\tau) = \langle \frac{\delta_i.\text{Sin}(\theta_i)}{T_{stop}(i)-\tau}, \frac{\delta_i.\text{Cos}(\theta_i)}{T_{stop}(i)-\tau} \rangle$.

Now we have,

$V5 = \{\langle \text{Occupies}(\text{Robot},\gamma),\tau\rangle \mid \gamma = \text{Displace}(\text{Disc}(0{\cdot}5),\text{LOC}(\tau))\}$.

V6 interprets HoldsAt for the fluents Facing and Moving, and is defined as follows.

$$V6 = \{\langle \text{Facing}(\theta),\tau\rangle \mid \theta = \text{RO}(\tau)\} \cup \{\langle \text{Moving},\tau\rangle \mid T_{go}(i) < \tau \leq T_{stop}(i) \text{ for some } i\}.$$

From V6, it can be verified that V satisfies Axiom (B4). V7 interprets HoldsAt for the fluents Blocked and Touching. Given V5, this can be directly filled in so as to satisfy Axioms (B4) and (B6). The details are omitted, but it can be confirmed from the definition of V5 that, for any $0 < i \leq m$, there is some ω such that,

$\langle \text{Blocked}(\text{Robot},\omega,\theta),\tau_i\rangle \in V7$

and that if, for any any τ, ω and θ,

$\langle \text{Blocked}(\text{Robot},\omega,\theta),\tau\rangle \in V7$

then either,

- $\tau = \tau_i$ for some $0 < i \leq m$, or
- $V \Vdash \neg \text{HoldsAt}(\text{Moving},\tau) \wedge \neg \text{Happens}(\text{Go},\tau)$, or
- $V \Vdash \neg \text{HoldsAt}(\text{Facing}(\theta,\tau)$.

Given this, it can be verified that V satisfies Axiom (H1) plus its only-if counterpart.

Similarly, it can be verified from the definition of V5 that V satisfies Axiom (H2) and (H3) plus their only-if counterparts. Since (H1) to (H3) are satisfied along with their only-if counterparts, V is minimal with respect to Happens, and V satisfies CIRC[N1 $\wedge$ N2 ; Happens].

Let Initiates, Terminates, and Releases be interpreted as follows.

$$V[\![\text{Initiates}]\!] = \{\langle \text{Go},\text{Moving},\tau\rangle\} \cup \{\langle \text{Rotate}(\theta_1),\text{Facing}(\theta_1+\theta_2),\tau\rangle \mid \theta_2 = \text{RO}(\tau)\} \cup \{\langle \alpha,\text{Occupies}(\text{Robot},\gamma),\tau\rangle \mid \alpha = \text{Stop or Bump and } \gamma = \text{Displace}(\text{Disc}(0{\cdot}5),\text{LOC}(\tau))\}$$

$$V[\![\text{Terminates}]\!] = \{\langle \text{Stop},\text{Moving},\tau\rangle\} \cup \{\langle \text{Bump},\text{Moving},\tau\rangle\} \cup \{\langle \text{Rotate}(\theta),\text{Moving},\tau\rangle\}$$

$V[\![\text{Releases}]\!] =$
$\{\langle \text{Go}, \text{Occupies}(\text{Robot},\gamma),\tau\rangle\} \cup$
$\{\langle \text{Rotate}(\theta_1), \text{Facing}(\theta_2),\tau\rangle \mid \theta_2 = \text{RO}(\tau) \text{ and } \theta_1 \neq 0\}$

It can easily be verified that V satisfies CIRC[E ; Initiates, Terminates, Releases], given Lemma B.1.

Now it's a routine exercise to fill in the interpretations of Clipped and Declipped according to Axioms (EC5) and (EC6). Finally, based on V5, an interpretation of the Trajectory predicate can be supplied which conforms to Axioms (B7) and (B8), and it can be verified that V satisfies the axioms of the event calculus. □

Proof of Theorem 8.10. Given that $O \wedge M1 \wedge M2$ is consistent, a model can be constructed for Σ as shown in the proof of Theorem 8.9. □

Proof of Theorem 9.3. Noting that no square in M can be both black and white, we can see that $O \wedge M1 \wedge \text{Exp}(M)$ is consistent.

It can be verified by inspection of the algorithm that the sequence of points P[1] to P[m] is a bump trail. Let ρ_i be the i^{th} element of P. We now proceed to show that, for any i, $0 < i \leq m$,

$$O \wedge M1 \wedge \text{Exp}(M) \vDash \text{OTR}(\rho_i, \text{BD}_{min}(\eta_i, \text{RO}(\tau_i)), \text{BD}_{max}(\eta_i, \text{RO}(\tau_i))). \qquad \text{[B.11]}$$

From Lines 11 and 15 of ColourMap and the definition of PossSquare, we know that M will include a black square with a boundary which, if it were the boundary of an object, would meet the conditions set out in the definition of OTR for a boundary l. From Line 12 of ColourMap and the definition of ColourPath, we see that this boundary will also be shared by a white square. So, according to the definition of Exp, it will indeed be the boundary of an object.

Next we show that,

$$\begin{aligned}&\text{CIRC}[O \wedge M1 \wedge \text{Exp}(M) ; \text{AbSpace} ; \text{Initially}_P, \text{Initially}_N] \wedge \\ &\quad \text{CIRC}[N1 \wedge N2 ; \text{Happens}] \wedge \\ &\qquad \text{CIRC}[E ; \text{Initiates}, \text{Terminates}, \text{Releases}] \wedge B \nvDash \neg \text{COMP}[\Psi].\end{aligned} \qquad \text{[B.12]}$$

From Lines 7 and 12 of ColourMap and the definitions of PlotPath and ColourPath, it can be seen that the white squares in M at the end of ColourPath's execution cover a clear path for the robot from its starting position to the site of its first change of direction, and from each point where it changes direction to the next. Since the corresponding region of space in Exp(M) will be empty, then models of,

$$\begin{aligned}&\text{CIRC}[O \wedge M1 \wedge \text{Exp}(M) ; \text{AbSpace} ; \text{Initially}_P, \text{Initially}_N] \wedge \\ &\quad \text{CIRC}[N1 \wedge N2 ; \text{Happens}] \wedge \\ &\qquad \text{CIRC}[E ; \text{Initiates}, \text{Terminates}, \text{Releases}] \wedge B\end{aligned}$$

exist in which the only collisions are those recorded in Ψ.

Given the consistency of $O \wedge M1 \wedge Exp(M)$, the theorem follows from [B.11], [B.12], and Theorem 8.9. □

Proof of Theorem 9.5. Since M2 is curve-free, for every bump trail ρ_1 to ρ_m such that, for any $0 < i \leq m$,

$$O \wedge M1 \wedge M2 \models OTR(\rho_i, BD_{min}(\eta_i, RO(\tau_i)), BD_{max}(\eta_i, RO(\tau_i)))$$

there exists some P[1] to P[m] which is a possible result of the algorithm's non-deterministic choices on Line 11 of ColourMap and Line 5 of PlotPath, such that, for any $0 < i \leq m$, $P[i] = \rho_i$. Therefore, from Theorem 8.8, the algorithm's non-deterministic choices cover all possible explanations of the requisite form. □