The Formalities of Affordance

Antony Galton[1]

Abstract. It is an obvious truth that the possibilities for action and movement are conditioned by the physical spatial environment. In the terminology of J. J. Gibson, these possibilities are defined by the "affordances" of environmental features, and the key to being a successful agent in the physical world is being able to perceive and exploit these affordances. To what extent can affordances be characterised in terms of low-level environmental features using the methods of traditional logic-based commonsense knowledge representation? Following an introductory general discussion, this paper concentrates on a particular case, the affordance of containment, for which we give a sequence of successively more detailed and lower-level analyses.

1 Affordances and Image Schemas

The notion of *affordance* was introduced by J. J. Gibson as part of his radical "ecological" theory of perception. Whereas previous theories had held that an individual's perception of its environment must be mediated by percepts corresponding to the ever-shifting patterns of sensory stimulation to which the individual is subject, Gibson believed that the environment is perceived *directly*, in the form of the ambient array of surfaces constituting the environment within which the individual moves and acts. Although the patterns of sensory stimulation must clearly play a part in giving rise to the direct perception of the surrounding surfaces, they are not themselves perceived, but serve merely as conduits by which the information contained in those surfaces is brought to the attention of the perceiver. That we are not aware of the patterns of sensory stimulation themselves should be sufficiently obvious if we consider the case of the eye: if we were able somehow to observe the patterns of light falling on the retina, we would certainly not be able to discern from these the external world which, in practice, we perceive with such immediacy; instead, all we should see would be a "blooming buzzing confusion", as a result of the rapid movements of the eyeballs as well as the movements of the subject's own head and body.

A key feature of Gibson's theory is the further observation that the potentialities for movement and action afforded to an individual by its environment are inherent properties of the surface layout by which the environment is defined. As Gibson himself puts it,

> Perhaps the composition and layout of surfaces *constitute* what they afford. If so, to perceive them is to perceive what they afford. This is a radical hypothesis, for it implies that the "values" and "meanings" of things in the environment can be directly perceived. Moreover, it would explain the sense in which values and meanings are external to the observer. [11, p.127]

Thus we perceive directly that a firm, more or less horizontal surface supported about 50cm above the surrounding ground is suitable, if sufficiently wide and deep, for sitting on — it "affords sitting" — and a sufficiently high and wide aperture in a more or less vertical solid surface can be passed through (it "affords entering"). A well-known attempt to make explicit the physical properties that a surface layout must exhibit in order to be possessed of a certain affordance is that of Warren [25] who, amongst other things, shows experimentally that in order for a set of stairs to be climbable for a given subject, the ratio between the vertical height of each step and the subject's own leg-length should be not more than 0.88.

In investigating affordances we should distinguish between several distinct goals, all of which must be achieved before a complete theory can be obtained. We may refer to these as upper-, middle- and lower-level goals, and they may be formulated as follows:

1. The upper-level goal is to answer what may be called "ecological" or "environmental" questions concerning the role of affordances in the life of an individual, how they can be used to explain features of human and animal behaviour, and how they can be exploited for the better design of environments.
2. The middle-level goal is concerned with characterising exactly what affordances are: this may be called the "ontological" question. How is an affordance defined, and what is the logical relationship between statements about affordances and other statements about the world?
3. The lower-level goal is the answer the "aetiological" question of where affordances come from, exactly how the physical layout of surfaces determines the affordances it has for any given class of creatures.

As an example of the middle-level goal, Steedman [22, 23] considers the affordances associated with doors. He uses a *linear dynamic event calculus* to formalise such statements as that a door can be gone through if open, but not if shut; if it is pushed when shut, it becomes open, and vice versa; if one is inside, then the result of going through the door is to be outside, and vice versa. These capture the affordances of a door *qua* passageway as well as *qua* barrier. On the other hand, no consideration is given to the physical characteristics that something must have in order to be able to function as (i.e., possess all the relevant affordances of) a door.

As pointed out by Frank and Raubal [7] and elaborated further by Kuhn [17], affordances are closely related to *image schemas* [18], recurring patterns which we employ to structure our understanding of the world we live in, and which are presumed to play a fundamental role in human cognition and language. Examples of image schemas include CONTAINER and PATH: the link with affordances is obvious, since to be a container is precisely to *afford containment*, while to be a path is to *afford passage*. Thus at least in many cases image schemas may be characterised in terms of the affordances of actual exemplars of those schemas.

An example of the upper-level goal is Jordan *et al*'s sketch for an affordance-based model of *place* in GIS [16]. As is well recognised, the notion of place is complex, not to be reduced to some simplistic

[1] University of Exeter, UK, email: apgalton@ex.ac.uk

construct in terms of location. A place is a portion of the environment that can fulfil certain purposes of an agent or community of agents, e.g., "here is a good place to have our picnic". In order for X to be a good place for Y to Z it is necessary, at least, that X affords Zing to Y. Jordan *et al* provide a useful discussion of the many factors that need to be taken into account in giving an affordance-based model of place, though their claim to have presented a "methodology to model places with affordances" is perhaps overstated.

This paper is concerned less with the upper or middle level goals than with the comparatively neglected lower level: in virtue of what does a given surface layout possess a particular set of affordances? Some of the work of Warren cited above (e.g., concerning the physical requirements for steps to be climbable) falls under this category. This particular example is quantitative in nature: the climbability of steps is referred to numerical measurements of both steps and climber. This is obviously important, since the numerical measurements make all the difference. Nonetheless, the quantitive questions cannot really be asked unless certain qualitative conditions are satisfied first: to be a candidate for being a flight of stairs, for example, there must exist an appropriately configured sequence of alternating horizontal surfaces and vertical displacements, and in the absence of this, or a close enough approximation, there is nothing to measure!

In the spirit of qualitative reasoning in AI, our aim is to consider the low-level question from a qualitative point of view. In particular, we shall be concerned with the following question: To what extent can the affordance-generating features of surface layouts be specified in terms of simple qualitative calculi such as RCC [20]? The analysis will be very much in the spirit of the formalisations of commonsense knowledge exemplified by such works as [12, 13, 21, 5], in which the notion of affordance is certainly frequently implicit, even if not brought to the fore as the explicit focus of attention. In the space available, it will only be possible to look in detail at one particular type of affordance, that of *containment*.

2 A Few Preliminaries

An affordance is a potentiality for action offered by some environmental feature to an agent. Gibson stressed the *mutuality* implicit in this definition — it takes two to make an affordance, that which affords something, and that to which it affords it. Thus the formal expression of affordance must be relational in nature. The "action" which is afforded does not necessarily involve motion (e.g., a text affords reading, a bed affords sleeping) but in the most typical cases it does so. Hence the formal expression of affordance will often involve an analysis of some kind of motion. For this, we require an appropriate grounding in spatial and temporal representations.

2.1 Spatial regions

In this paper we use the well-known RCC system of [20], and specifically the following relations:

$P(x,y)$	x is part of y
$PP(x,y)$	x is a proper part of y
$TP(x,y)$	x is a tangential part of y
$TPP(x,y)$	x is a tangential proper part of y
$EC(x,y)$	x is externally connected to y
$O(x,y)$	x overlaps y
$PO(x,y)$	x partially overlaps y

We treat these as relations between spatial regions rather than objects (but see below, §2.2). Spatial relations between objects are expressed using RCC relations between their *positions*. The position of object o at time t, denoted $pos(o,t)$, is a spatial region which coincides with the spatial extent of o at t. Note that pos encodes both position and shape: if o changes shape, then $pos(o,t)$ changes. Use of this notation does not presuppose a "Newtonian" notion of absolute space: as in everyday life, positions are always specified in a relative way, by reference to some framework of objects which, for the purposes at hand, can be regarded as fixed (e.g., the walls of a room, the surface of the earth) even though from some wider perspective they may be regarded as moving. The use of pos only presupposes that we have some such framework implicitly to hand. For a detailed discussion of the related notion of "relative place", see [6].

We shall not attempt to define exactly what a spatial region is, but merely content ourselves with the observation that a spatial region is a possible position for an object. As such, a region is paradigmatically three-dimensional, since material objects are. However, we will also need to refer to the boundary (or surface) of a three-dimensional region, and this is of course two-dimensional. We write ∂x to denote the boundary of x. Other spatial notions needed are the *convex hull* of a region r, denoted $cvhull(r)$, and the relation of *congruence* between spatial regions, denoted $Congruent(r_1, r_2)$. This must be stipulated to be an equivalence relation, and in addition it should satisfy the rule that any part of a region congruent to a given region z must be congruent to part of z, i.e.,

$$P(x,y) \wedge Congruent(y,z) \rightarrow \exists u(Congruent(x,u) \wedge P(u,z)) \quad (1)$$

but no further attempt will be made to define it here.

Although we make no commitment to a representation of regions as sets of points, we will make use of the set-theoretic notations for union ($\cup$), intersection ($\cap$), and difference ($\setminus$) on the assumption that some suitable analogous operations would be definable in any "pointless" theory of regions.

2.2 Physical objects

Physical objects include both *material* objects, which are made of matter, and *non-material* objects such as holes (and in particular the inner spaces of containers) which are dependent on material objects but not themselves material. Both material and non-material pbjcts have positions. The predicate $Material(x)$ is used to say that an object is material.

We shall use a parallel series of RCC relations, notated P^*, PP^*, TP^*, etc, to apply to physical objects, where "connection" is now understood to mean physical attachment rather than spatial contiguity. Thus objects are EC^* if they are actually joined together, DC^* if not. If objects o_1 and o_2 are not joined together but are touching at time t, the relation between them can be expressed as $DC^*(o_1, o_2) \wedge EC(pos(o_1,t), pos(o_2,t))$.

Material objects are characterised by the *principle of non-interpenetrability*, which says that non-overlapping material objects cannot simultaneously occupy overlapping positions, i.e.,

$$Material(o_1) \wedge Material(o_2) \wedge \neg O^*(o_1, o_2) \rightarrow \forall t \neg O(pos(o_1,t), pos(o_2,t)) \quad (2)$$

2.3 Time

We shall have cause to refer to specific motion events. A number of different formalisms are available for this purpose, notably the method of temporal arguments, event-type reification, and event-token reification; these methods are described and compared in [10]. In this paper, we use the method of temporal arguments, by which, to say that an event of type E occurs over an interval $[t_1, t_2]$, we write $E(t_1, t_2)$, using the terms t_1 and t_2 as temporal arguments to the predicate E expressing the event type. Conversion to the other formalisms is mostly straightforward, if it is desired to go on to exploit the greater expressivity of those formalisms.

2.4 Modality and possible futures

Since it refers to what *can* happen rather than to what *does* happen, affordance is a *modal* notion. Its formal expression must therefore use either modal logic or some other formalism capable of expressing an appropriate notion of modality. It is a non-trivial task to specify exactly the notion of modality we require, and some discussion of this is needed before we can proceed further.

Affordances are important because of their role in determining possible future actions: affordance is a potentiality, and what is now afforded, and therefore potential, may become actualised in the future. If we talk about the affordance that something had at some past time, we are implicitly referring to the possible futures running forward from that time. Thus the form of modality appropriate for describing affordances is future-directed: with reference to any time point, we are interested in its different possible futures, but regard its past as fixed. The possibility operator we will use may be characterised informally as follows: $\Diamond P$ is true at t if and only if there is some possible future of t such that, if that future is the actual future, then P is true at t. This can be described formally in terms of the *history structures* of [1], in which this operator is notated $\circledast$.[2] Under this interpretation, the logic of $\Diamond$ is given by the modal system S5.

This does not, however, fully characterise the meaning of the operator. What does "possible future" mean? To illustrate the problem, consider a slot machine which will accept 1 euro coins: this means (at least) that the slot affords entry by a 1 euro coin. Does the slot afford entry by a metal sphere of diameter 12mm? The practical commonsense answer is "no": the sphere is too wide to fit into the slot. But what if I take a hammer and flatten the sphere into a disk? Then it will surely fit into the slot (the volume of the sphere is a little less than that of a 1 euro coin). The natural reply to this is to say that this is "cheating": it is not what we meant when we asked if the slot could admit the metal sphere. But now consider this case: I have written a letter on an A4 sheet of paper; I have an A5-sized envelope: can I use it to post my letter? This time the answer is surely "yes": I can fold my letter in two and slip it into the envelope. The envelope affords containment for an A4 sheet. Somewhere between folding a sheet of paper in two and hammering a metal sphere flat lies the borderline between those possible histories which we wish to allow for the purpose of defining affordances and those which we do not. But where exactly? The paper-folding is more easily reversible than the sphere-flattening: but it is not completely reversible, since you can never get rid of a crease in a sheet of paper. Perhaps the key lies in the notion that folding a letter to fit it into an envelope is an entirely normal and expected procedure: it is what we do. Flattening a sphere to fit it into a slot intended for coins is highly unusual and only likely to be done under exceptional circumstances.

We therefore do not want our modal operators to range over all conceivable futures, or even all physically possible ones. Somehow we must restrict our attention to those futures in which exceptional, affordance-disrupting events do not occur except perhaps in exceptional circumstances. It is in terms of these futures that the modal operators $\Box$ and $\Diamond$ are to be interpreted. Of course, to say this is to say virtually nothing until we have characterised what "normal" or "expected" means. We acknowledge the ultimate necessity of doing this, but meanwhile proceed to the technicalities of characterising affordances on the assumption that some suitable definition of the modal operators can be given.

[2] This operator is also used — notated M — in [8, Ch.7].

2.5 Rigidity

Modality enters into the definition of another physical property, namely *rigidity*, which will be important in what follows. A material object is rigid if it *cannot* change shape. In reality, of course, absolute rigidity is a fiction, but in practice many objects can be treated as if they were rigid, and in particular for the logic of containment the distinction between rigid objects such as apples and boxes and non-rigid objects such as bags and scarves (and human bodies!) is important. An object is rigid if all of its possible positions are congruent:

$$\begin{aligned}&Rigid(o) \ =_{\mathrm{df}}\\ &\forall t \forall t' \forall r \forall r' (\Diamond(pos(o,t) = r) \wedge \Diamond(pos(o,t') = r') \ \rightarrow \\ &\qquad\qquad\qquad\qquad Congruent(r,r'))\end{aligned} \tag{3}$$

3 Case study: The Affordance of Containment

What is a container? It is hard to give a non-circular answer. A container is something which can contain other things. What does it mean to contain something? For A to contain B is for B to be in A. What does "in" mean? "A is in B" means that B contains A ... and we are back where we started.

We might say that for A to contain B is for A to constrain the position of B in a certain way. For example, the coins in my pocket go wherever my pocket goes, unless they are taken out. The water in a jug is held in place by the jug — without the jug, the water would spread out and find its way to the lowest accessible spaces. But if a man is in a house, in what sense is his position constrained by the house? What about a tree in a field?

3.1 Contained space

We do not attempt to *define* containers or the containment relation here; but we can at least try to say as much as we can about it that is clear, definite, and formalisable. To this end, we make use of the notion of the *contained space* of a container, introduced by Hayes [13] in the context of containers for liquids. A cup, for example, is a solid ceramic object used for containing liquids; its contained space is the space partially enclosed by the material of the cup, within which anything contained by the cup is located. As Hayes says, the contained space "is *not* a physical object but is characterized by a certain capacity and by being in a certain relation to a container". In our terminology, it is a non-material object dependent on the container. The contained space of a container is well-defined since there is a point beyond which, if more liquid is added, it will overflow; the surface of the liquid at this point defines the upper boundary of the contained space. For solid or granular matter, it is harder to specify that part of the boundary of the contained space which is not shared with the container itself. We will not attempt to address this problem here.[3]

[3] Note also that for solids, we often speak of containment even when only part of the contained object is actually in the contained space of the container, e.g., a vase containing flowers, where only the stalks of the flowers are actually inside the vase. Such examples have been discussed extensively in the literature on spatial prepositions [15, 24, 3].

The contained space $cs(x)$ of a container x belongs to x, but is not part of it. It is not a spatial region, since it may be located at different spatial regions at different times; as x moves around, $cs(x)$ moves with it.[4] In particular, $cs(x)$ is always contiguous with x, i.e.,

$$Container(x) \rightarrow \forall t EC^*(cs(x), x).$$

and it is always located within the convex hull of the region occupied by x:

$$Container(x) \rightarrow \forall t(P(pos(cs(x), t), cvhull(pos(x), t))).$$

A container is *closed* when the boundary of its contained space forms part of the boundary of the container itself:

$$Closed(x,t) =_{\mathrm{df}} Container(x) \wedge P(\partial pos(cs(x), t), \partial pos(x, t)).$$

When a container is not closed, the connected components of $\partial pos(cs(x), t) \setminus \partial pos(x, t)$ are (again following [13]) called *portals* of the container. They are the entrances and exits by which objects can enter or leave the container. Many containers exhibit both open and closed states: a box with a hinged lid, for example, is open when the lid is raised, and closed when it is down.

3.2 Containment

To say that container c contains object o at a given time is to say that the spatial region occupied by o at that time is part of the spatial region occupied by the contained space of c, i.e., [5]

$$Contains(c, o, t) =_{\mathrm{df}} P(pos(o, t), pos(cs(c), t)) \tag{4}$$

To formalise the notion that c affords containment for o, we need to say that c *can* contain o. At time t, this will be the case so long as it is possible for c to contain o at some time at or in the future of t:

$$CanContain(c, o, t) =_{\mathrm{df}} \exists t'(t \leq t' \wedge \Diamond Contains(c, o, t')). \tag{5}$$

It is implicit in the use of $\Diamond$ here that o can be moved over to c and can enter it, with either or both undergoing changes of shape needed to allow this to happen — all such changes being of the kind we have called "normal" or "expected".

3.3 Containment and rigidity

In general, we do not wish to make any assumptions concerning the rigidity or otherwise of containers and what can be put in them. We have four distinct cases here, as shown in the following table:

	Rigid container	Non-rigid container
Rigid object	Apple in a box	Apple in a bag
Non-rigid object	Scarf in a box	Scarf in a bag

We consider the form taken by $CanContain$ in the case where c and o are both rigid. It seems obvious that if a container is rigid, then so is its contained space; so we postulate the rule

$$Container(c) \wedge Rigid(c) \rightarrow Rigid(cs(c)) \tag{6}$$

[4] The ontology of contained spaces is similar to the ontology of holes [2], sharing many of the same problems and difficulties.

[5] We do not need to specify that c is a container; if c is not a container, $cs(c)$ can be defined to be the null object, so $pos(cs(c), t)$ will also be null, and hence $Contains(c, o, t)$ is necessarily false.

Hence $cs(c)$ is rigid as well. Assume $CanContain(c, o, t)$. From (5) this means there is a time t' not earlier than t such that $\Diamond Contains(c, o, t')$, i.e., from (4),

$$\Diamond P(pos(o, t'), pos(cs(c), t')).$$

What regions exist cannot vary between different possible futures, so this implies there are regions r_1 and r_2 such that

$$P(r_1, r_2) \wedge \Diamond(pos(o, t') = r_1) \wedge \Diamond(pos(cs(c), t') = r_2).$$

By (3), since o and $cs(c)$ are both rigid, this implies[6]

$$P(r_1, r_2) \wedge Congruent(r_1, pos(o, t)) \wedge Congruent(r_2, pos(cs(c), t)).$$

Finally, from (1), since congruence is an equivalence relation, we obtain $\exists u(Congruent(pos(o, t), u) \wedge P(u, pos(cs(c), t)))$. Thus, a rigid container can contain a rigid object only if the latter is congruent to part of the contained space of the latter:

$$CanContain(c, o, t) \wedge Rigid(c) \wedge Rigid(o) \rightarrow \exists u(Congruent(pos(o, t), u) \wedge P(u, pos(cs(c), t)))$$

3.4 Trajectories and continuity

Although we have seen how to say that a container can contain an object, we have not really addressed our original question, which is *by virtue of what* does this potentiality obtain? Consider the case in which object o is outside container c at time t_0, and inside it at later time t_1. Over the interval $[t_0, t_1]$ both o and c may change position and shape. We can track the values of $pos(o, t)$ and $pos(c, t)$ as t runs from t_0 through to t_1. These specify the *trajectories* of o and c. A condition for an object to come to be inside a container is that suitable trajectories for both the object and the container exist.

A *trajectory* is simply a continuous sequence of spatial regions. Formally, it may be represented by a continuous function

$$traj : [0, 1] \rightarrow \mathcal{R},$$

where $\mathcal{R}$ is the set of all spatial regions.[7] Exactly what is meant by "continuous" here needs discussion. A number of approaches to this have been suggested in the literature. One way is to adopt a four-dimensional view, and try to characterise continuity in terms of the shape of the spatio-temporal extent of the motion considered as a region in four dimensions [19, 14]. Another approach, closer in spirit to our current enterprise, is to characterise continuity in terms of some metric on the space of possible regions [9, Ch. 7][4, 5]; metrics considered include the Hausdorff distance and variations on that, and the volume of the symmetric difference. For each such metric Δ, a trajectory $traj$ may be characterised as continuous with respect to that metric in the usual way, i.e.,

$$\forall t \in [0, 1] \forall \epsilon > 0 \exists \delta > 0 \forall t' \in [0, 1](|t - t'| < \delta \rightarrow |traj(t) - traj(t')| < \epsilon).$$

[6] Here we are also using the trivial facts $\Diamond(pos(o, t) = pos(o, t))$ and $\Diamond(pos(cs(c), t) = pos(cs(c), t))$.

[7] Expressions of the form $traj(x)$ should be understood as "syntactic sugar" for something along the lines of $value(traj, x)$, where $traj$ is a term rather than a function symbol; this will allow us to quantify over trajectories without breaking the bounds of first-order logic. But for ease of reading we shall retain the notation $traj(x)$.

The particular trajectories we are interested in are sequences of possible positions of an object. The following formula says that o follows trajectory $traj$ over the interval $[t_0, t_1]$:

$$\begin{aligned} &Follows(o, traj, t_0, t_1) =_{\mathrm{df}} \\ &\qquad \forall t \left(t_0 \leq t \leq t_1 \rightarrow pos(o,t) = traj\left(\tfrac{t-t_0}{t_1-t_0}\right)\right) \end{aligned}$$

Of course, this trajectory cannot be followed if there are obstacles in the way; but this need not be specified explicitly, since given $Follows(o, traj, t_0, t_1)$, non-interpenetrability (2) already implies that when o is at any point in the trajectory, no other body overlaps the position it then occupies. It is not necessary for the whole trajectory to be unoccupied by other objects throughout $[t_0, t_1]$: an obstacle is fine so long as it is removed when you get to it.

Continuity of motion is now secured by means of the rule

$$\begin{aligned} &\forall t \forall t' \forall r \forall r' (t < t' \wedge pos(o,t) = r \wedge pos(o,t') = r' \rightarrow \\ &\quad \exists traj(traj(0) = r \wedge traj(1) = r' \wedge \\ &\qquad Follows(o, traj, t_0, t_1))) \end{aligned} \tag{7}$$

3.5 Entry into a container

In order for o to come to be inside c during the interval $[t_0, t_1]$, o and c must follow trajectories which begin with o and c in positions such that o is outside c, and end with them in positions such that c contains o, and which are such that at no time do the positions of o and c overlap. This motion can be divided into three parts: first, o and c get into a position where o is "just outside" c; then o actually enters c, arriving at a position where it is "just inside" c; and finally, it may proceed to a resting position inside c. The middle phase is the crucial one: this is what we will call the *entering* event.

We must now formally characterise the actual process of entering. This begins at the latest time when o is "just outside" c, i.e., $EC(pos(o,t), pos(cs(c), t))$, and ends at the first time when o is "just inside" c, i.e., $TPP(pos(o,t), pos(cs(c), t))$. We can put

$$\begin{aligned} &Enters(o, c, t_0, t_1) =_{\mathrm{df}} \\ &\quad \exists traj_o \exists traj_c (\\ &\qquad Follows(o, traj_o, t_0, t_1) \wedge \\ &\qquad Follows(c, traj_c, t_0, t_1) \wedge \\ &\qquad \forall t (t_0 \leq t \leq t_1 \rightarrow \\ &\qquad\quad EC(pos(o,t), pos(cs(c), t)) \leftrightarrow t = t_0 \wedge \\ &\qquad\quad TPP(pos(o,t), pos(cs(c), t)) \leftrightarrow t = t_1)) \end{aligned} \tag{8}$$

It follows from non-interpenetrability (2) that the positions of o and c never overlap, i.e., we have $\neg O(pos(o,t), pos(c,t)))$ at all times t. Hence we do not need to specify this explicitly in (8).

For $t_0 < t < t_1$, we have $PO(pos(o,t), pos(cs(c), t))$, and if we assume that o is a one-piece object, this means that o must intersect a portal p of c. We will return to the implications of this later.

It should be emphasised that in order to get o inside c, either or both bodies may need to change shape (cf. the table above). This is allowed for in (8), since there is no reason why the values of $traj_o$ should all be congruent, and similarly for the values of $traj_c$. In particular, c may be closed initially; but this does not matter so long as a portal has opened at the time o needs to enter it.

From continuity, it seems plausible that

$$\begin{aligned} &\neg O(pos(o,t), pos(cs(c), t)) \wedge Contains(c, o, t') \rightarrow \\ &\quad \exists t_0 \exists t_1 (t \leq t_0 < t_1 \leq t' \wedge Enters(o, c, t_0, t_1)) \end{aligned} \tag{9}$$

Can we prove this from the rules and definitions we have given so far? If not, what further rules are needed? These are currently unanswered questions.

The definition (8) tells us what it is for o to enter c; but if we are interested in affordances, we want rather to specify what it means to say that o *can* enter c. The obvious definition is

$$CanEnter(o, c, t) =_{\mathrm{df}} \exists t' \Diamond Enters(o, c, t, t') \tag{10}$$

In particular, we would like to prove the modalised version of (9):

$$\begin{aligned} &\neg O(pos(o,t), pos(cs(c), t)) \rightarrow \\ &\quad (CanContain(o, c, t) \leftrightarrow \\ &\qquad \exists t' (t \leq t' \wedge CanEnter(o, c, t'))). \end{aligned} \tag{11}$$

This deceptively simple formula is none the less highly significant. The predicate $CanContain$ expresses the bare affordance of containment, which (4) defines as the potentiality for actual containment. This is, as we noted, a rather high-level view of the affordance, abstracting away from the features of the world in virtue of which containment is afforded in any particular situation. By invoking the principle of continuity of movement, we were able to express a more detailed precondition for the affordance of containment, namely, not just that an object can be situated inside a container, but that it can come to be there, in other words that there is a trajectory by which it can enter it. This is what is expressed by $CanEnter$; it gives us a somewhat lower-level view of the affordance. Formula (11) links these two views, the higher-level to the lower, by asserting in effect that they refer to the same underlying reality.

3.6 Entry at a portal

A still lower-level view is possible. Let us return to the observation that during the entering event, o must intersect a portal of c. Consider informally the preconditions for o to be able to enter c via the portal p. One is that c can contain o, i.e., its interior can be so located that o lies wholly inside it. Another is that there is a continuous series of cross-sections of o, each of which fits inside p. Here p and the cross-sections of o are two-dimensional entities. These conditions are not so far sufficient, as can be seen from Figure 1, where the vase is large enough to contain the ball, and every cross-section of the ball fits into the entrance portal of the vase (shown by the dotted line), but still the ball cannot enter the vase assuming both are rigid. (Of course a rubber ball could be squeezed past the constriction in the neck of the vase.) We need an additional condition, that there is a possible position of o inside c that is tangential to p.

How do we express these conditions formally? We assume here that o is a one-piece object (as opposed to, for instance, a two-piece object such as a teapot which consists of a body and a separate lid). Then $pos(o,t)$ is always a connected spatial region, which means that any two points within it can be joined by a one-dimensional

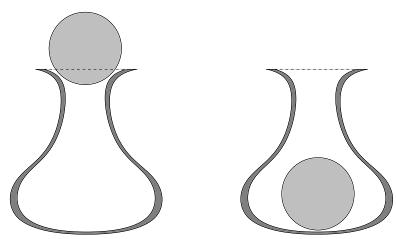

Figure 1. The vase could contain the ball if only the ball could get into it.

path lying wholly within the region. If $PO(pos(o,t), pos(cs(c),t))$, then from the definition of PO, part of o lies outside $cs(c)$ and part of o lies inside $cs(c)$. For each pair of points in $pos(o,t)$, one of which lies outside $cs(c)$ and the other inside, there is a path between them which (i) lies wholly within $pos(o,t)$ and (ii) intersects the boundary $\partial pos(cs(c),t)$. Let x be the sum of all the intersections of such paths with $\partial pos(cs(c),t)$, so we have $P(x, pos(o,t)) \wedge P(x, \partial pos(cs(c),t))$. By non-interpenetrability (2), o does not overlap c, and hence no part of o overlaps the boundary of c. In particular, $\neg O(x, \partial pos(c,t))$. This means that we have $P(x, \partial pos(cs(c),t) \setminus \partial pos(c,t))$, i.e., x is part of the portals of c.

By formalisation of this argument, one might hope to prove, from the principles enunciated so far, the following formula:

$$\begin{aligned}&Enters(o,c,t_0,t_1) \rightarrow\\ &\quad \forall t(t_0 < t < t_1 \rightarrow\\ &\qquad \exists r_1 \exists r_2 (pos(o,t) = r_1 \cup r_2 \wedge\\ &\qquad\qquad \neg O(r_1, pos(cs(c),t)) \wedge\\ &\qquad\qquad P(r_2, pos(cs(c),t)) \wedge\\ &\qquad\qquad P(\partial r_1 \cap \partial r_2, \partial pos(cs(c),t) \setminus \partial pos(c,t))))\end{aligned} \tag{12}$$

Note that this does not imply that $\partial r_1 \cap \partial r_2$ (our earlier x) is part of just one portal of c. A horseshoe-shaped object could have its two ends inserted into different entrances of a container with more than one entrance; but if the object is to enter the container, then all of it must pass through just one portal eventually. No doubt an argument based on continuity should enable us to establish this, but again, the details are at present unclear.

4 Concluding remarks

To summarise what we have done, we began with the goal of characterising in formal terms the conditions under which it can be said that a certain affordance exists, namely the affordance of containment which a container has in relation to an object. We began with a very high level characterisation which amounted to little more than a definition of what it means to say that one thing can contain another. This was the definition (5). By combining this definition with the condition of continuity (7), we were able to spell out a lower-level condition for the container to afford containment to the object, namely that it is possible for the object to enter the container; this is expressed in the definition of $CanEnter$ (10), the full details of which are contained in the definition of the $Enters$ predicate (8). Finally, by invoking the principle of non-interpenetrability (2), we were able, at least informally, to tease out a still lower-level condition for the affordance, relating the portals of the container to the sequence of cross-sections of the object which must intersect the portal as it enters the container. This was expressed, in part, by (12).

The general approach may be summarised as follows. To define what it means for some object or collection of objects to afford some action A to an object o, we begin by defining what it means for o actually to perform A, and then use a modalised form of this definition to provide a high-level definition of the affordance itself. Then, by invoking general principles such as continuity and non-interpenetrability, we tease out successively lower-level conditions for the affordance to exist. In this way, we gradually approach the goal of specifying just what it is about any particular physical layout that results in its having the affordances that it does. To relate this back to the original source of the affordance idea in Gibson's theories of perception, we can now say that we are able to perceive affordances by perceiving these lower-level conditions, which, we must assume, are more directly accessible to our perceptual apparatus.

It has to be admitted that so far much of this is programmatic. Even to handle fully the one case considered in this paper, namely containment, requires more detailed formal work than it has been possible to present here. Then there is whole field of enquiry ripe for investigation: affordance of shifting, lifting, hiding, opening, closing, and all the other potentialities offered by our environment which define the scope and limits of human action in the world.

REFERENCES

[1] B. Bennett and A. P. Galton, 'A unifying semantics for time and events', *Artificial Intelligence*, **153**, 13–48, (2004).

[2] Roberto Casati and Achille Varzi, *Holes and Other Superficialities*, MIT Press, Cambridge, MA and London, 1994.

[3] K. Coventry and S. Garrod, *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions*, Psychology Press, 2004.

[4] Ernest Davis, 'Continuous shape transformations and metrics of regions', *Fundamenta Informaticae*, **46**, 31–54, (2001).

[5] Ernest Davis, 'Pouring liquids: A study in commonsense physical reasoning', *Artificial Intelligence*, **172**, 1540–1578, (2008).

[6] M. Donnelly, 'Relative places', *Applied Ontology*, **1**(1), 55–75, (2005).

[7] Andrew U. Frank and Martin Raubal, 'Formal specification of image schemata — a step towards interoperability of geographic information systems', *Spatial Cognition and Computation*, **1**(1), 67–101, (1999).

[8] A. P. Galton, *The Logic of Aspect: an Axiomatic Approach*, Clarendon Press, Oxford, 1984.

[9] A. P. Galton, *Qualitative Spatial Change*, Oxford University Press, Oxford, 2000.

[10] A. P. Galton, 'Operators vs arguments: The ins and outs of reification', *Synthese*, **150**, 415–441, (2006).

[11] James J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, 1979.

[12] Patrick J. Hayes, 'The naïve physics manifesto', in *Expert Systems in the Micro-Electronic Age*, ed., D. Michie, 242–70, Edinburgh University Press, (1979).

[13] Patrick J. Hayes, 'Naive physics I: Ontology for liquids', in *Formal Theories of the Commonsense World*, eds., J. R. Hobbs and R. C. Moore, 71–107, Ablex, Norwood, New Jersey, (1985).

[14] S. M. Hazarika and A. G. Cohn, 'Qualitative spatio-temporal continuity', in *Spatial Information Theory: Foundations of Geographic Information Science*, ed., D. R. Montello, 92–107, Springer, Berlin, (2001).

[15] Annette Herskovits, *Language and Spatial Cognition*, Cambridge University Press, Cambridge, 1986.

[16] T. Jordan, M. Raubal, B. Gartrell, and M. J. Egenhofer, 'An affordance-based model of place in GIS', in *Eighth International Symposium on Spatial Data Handling, Vancouver, Canada*, eds., T. Poiker and N. Chrisman, pp. 98–109, (1998).

[17] Werner Kuhn, 'An image-schematic account of spatial categories', in *Spatial Information Theory, Proceedings of the 8th International Conference, COSIT 2007, Melbourne, Australia*, eds., S. Winter, M. Duckham, L. Kulik, and B. Kuipers, pp. 152–168. Springer, (2007).

[18] G. Lakoff and M. Johnson, *Metaphors We Live By*, University of Chicago Press, Chicago, 1980.

[19] Philippe Muller, 'A qualitative theory of motion based on spatio-temporal primitives', in *Principles of Knowledge Representation and Reasoning: Proceedings of KR'98*, eds., A.G. Cohn, L.K. Schubert, and S.C. Shapiro, San Mateo, Calif., (1998). Morgan Kaufmann.

[20] D. A. Randell, Z. Cui, and A. G. Cohn, 'A spatial logic based on regions and connection', in *Principles of Knowledge Representation and Reasoning: Proceedings of KR'92*, eds., B. Nebel, C. Rich, and W. Swartout, pp. 165–76, San Mateo, CA, (1992). Morgan Kaufmann.

[21] Murray Shanahan, 'An attempt to formalise a non-trivial benchmark problem in common sense reasoning', *Artificial Intelligence*, **153**, 141–165, (2004).

[22] Mark Steedman, 'Formalizing affordance', in *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*, pp. 834–839, (2002).

[23] Mark Steedman, 'Plans, affordances, and combinatory grammar', *Linguistics and Philosophy*, **25**(5), 723–753, (2002).

[24] Claude Vandeloise, *L'Espace en français*, Editions du Seuil, Paris, 1986.

[25] W. H. Warren, 'Constructing an econiche', in *The Ecology of Human-Machine Systems*, eds., J. Flach, P. Hancock, J. Caird, and K. Vicente, 210–237, Erlbaum, Hilldsale, NJ, (1995).