Qualitative Spatio-Temporal Representation and Reasoning:

Trends and Future Directions

Shyamanta M. Hazarika
Tezpur University, India

Information Science REFERENCE

Managing Director: Lindsay Johnston
Senior Editorial Director: Heather A. Probst
Book Production Manager: Sean Woznicki
Development Manager: Joel Gamon
Acquisitions Editor: Erika Gallagher
Cover Design: Nick Newcomer

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Qualitative spatio-temporal representation and reasoning: trends and future directions / Shyamanta M. Hazarika, editor.
p. cm.
Includes bibliographical references and index.
Summary: "This book is a contribution to the emerging discipline of qualitative spatial information theory within artificial intelligence, covering both theory and application-centric research and providing a comprehensive perspective on the emerging area of qualitative spatio-temporal representation and reasoning"-- Provided by publisher.
ISBN 978-1-61692-868-1 (hardcover) -- ISBN 978-1-61692-870-4 (ebook) -- ISBN 978-1-4666-1654-7 (print & perpetual access) 1. Qualitative reasoning. 2. Spatial analysis (Statistics) 3. Space and time--Mathematical models. 4. Logic, Symbolic and mathematical. I. Hazarika, Shyamanta M.
Q339.25.Q84 2012
006.3'33--dc23
2011050504

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 3
Qualitative Reasoning and Spatio-Temporal Continuity

Ernest Davis
New York University, USA

ABSTRACT

This chapter discusses the use of transition graphs for reasoning about continuous spatial change over time. The chapter first presents a general definition of a transition graph for a partition of a topological space. Then it defines the path-connected and the homogeneous refinements of such a partition. The qualitative behavior of paths through the space corresponds to the structure of paths through the associated transition graphs, and of associated interval label sequences, and the authors prove a number of metalogical theorems that characterize these correspondences in terms of the expressivity of associated first-order languages. They then turn to specific real-world problems and show how this theory can be applied to domains such as rigid objects, strings, and liquids.

1. INTRODUCTION

Many spatial aspects of many persistent entities vary continuously over time: the direction of a weather vane, the length of a rubber band; the shape of a balloon and so on. Many others, of course, do not: the territory of the United States, the shape of a shadow on a surface, the shape of a tree when a limb is pruned. However, when it is known that a spatial entity does change continuously, that constraint can be very useful in reasoning about its behavior over time.

Consider the following inferences:

A. Two interlocked jigsaw puzzle pieces cannot be separated by a movement in the plane of the puzzle, but can be separated by lifting one perpendicular to the plane.
B. Consider a string loop of length *L* wrapped once around the waist of an hourglass with

DOI: 10.4018/978-1-61692-868-1.ch003

spherical globes of circumference C. If $L>C$ then the loop can be removed from the hourglass without coming into contact with the hourglass and without ever being taut. If $L<C$, then the loop cannot be removed from the hourglass. If $L=C$, then the loop can be removed from the hourglass, but at some point it must be in contact with the hourglass, and it must be taut. It can be taken off either the upper or the lower globe.

If the globes of the hourglass are long cylinders, the circular cross section has circumference C, and $C=L$, then the string can be removed from the hourglass, but it will be taut and in contact with the hourglass over an extended interval of time.

If, instead of a string loop, we have a rubber band whose length is less than C at rest but can be stretched to a length greater than C, then it can be removed from the hourglass without being in contact with the hourglass, but it must be stretched in order to do so.

C. A quantity of milk in a closed bottle remains in the bottle. If at time $T1$ there is milk sitting in cup A, and at a later time $T2$ this milk has moved to a cup B, and both cups are stationary, then the milk came out of the top of cup A and went in the top of cup B.

D. The dog can go from the dining room into the kitchen. However, if a chair is placed in the middle of the kitchen doorway, then the dog cannot go from the dining room to the kitchen. If the chair is placed at the edge of the doorway, then the dog can squeeze past and get into the kitchen.

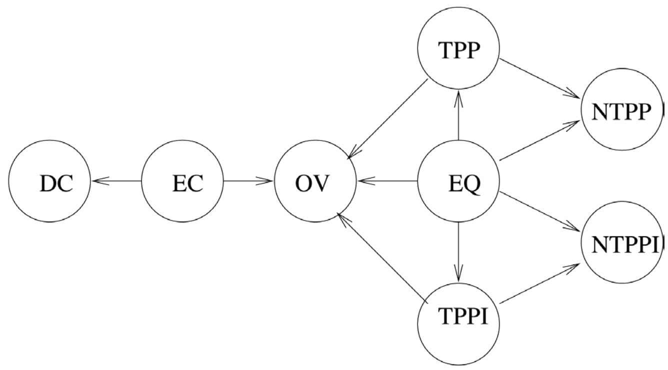
E. A person who is in Canada at one time and in the United States at a later time must cross the U.S. border at some time in between. A person who is in Alaska at one time and in Idaho at a later time must cross the U.S. border at least twice in between. It is possible to travel from any point in Idaho to any point in Ohio without crossing the border of the United States. This seems like the simplest of these inferences; in fact, however, it is the example for which the theory we develop in this chapter is least adequate.

A number of points may be observed about these examples. First, both the givens and the conclusions are qualitative; no precise measurements or shape descriptions are given. Second, they depend on continuity: If Star Trek style teleportation were available, the inferences would not be valid. For that matter, analogous inference can fail if they involved entities that change discontinuously; for instance, when the Louisiana Purchase took effect in 1803, many objects went from being far from the United States to being deep inside the United States without ever being on the border. Third, many of the sample inferences above depend on further physical limitations on the dynamic spatial behaviors of the objects involved in addition to continuity.

Moreover, the well-known scheme for representing qualitative spatial change in terms of transitions between RCC relations is inadequate to justify or represent these inferences. In that theory, as we will discuss in greater detail in section 2, the relation between two regions is characterized in terms of eight possible mereotopological relations. Spatial change over time is characterized in terms of the sequence of the evolving sequence of these relations. Continuity is characterized in terms of possible and impossible transitions from one relation to another, as illustrated in the well known graph in Figure 1.

However, this representation is not expressive enough to deal with examples like those above, for a number of reasons. First, the set of states corresponding to a particular RCC relation is sometimes *disconnected* and we sometimes wish to distinguish different connected components. For instance, in example (A) we wish to distinguish the states where the jigsaw puzzle pieces are Externally Connected (EC) and interlocked from those where they are externally connected and

Figure 1. Transition graph for RCC relations



separable. Second, the set of states corresponding to a particular RCC relation is sometimes *non-uniform* with respect to the possible transitions. For instance, in example (B), there is a transition from the contact state to the non-contact state, but that transition cannot be executed immediately when the string is around the middle of the cylinder. Third, we sometimes wish to make qualitative distinctions other than the RCC relations. For instance, in example (B), we wish to distinguish states in which the rubber band is *relaxed* from those in which it is *stretched*.

The aim of this chapter is to present a generalization of the RCC transition graph that addresses these issues for a broad range of constraints on shapes, definitions of continuity, and qualitative relations. The essential idea is that, given a space of object configurations and a partition of that space into qualitative categories, we further subdivide each category into cells that are both connected and uniform with respect to transitions. That is, any two configurations in the same cell are connected by a path that remains within the cell and have the same possible transitions to other cells. This is called the *Qualitative Homogeneous Decomposition* (QHD) of the starting partition (section 3). The QHD can be viewed as a directed graph, called the QHD graph (section 3.1). We give a precise meta-logical characterization of the expressivity of QHD graphs as encodings of qualitative changes in dynamic systems (section 3.2). We then discuss the ways in which this representation can be applied to specific examples like those above (section 4).

2. RELATED WORK

Work on both qualitative and precise spatial reasoning using continuity constraints, including certainly the current chapter, has mostly been centered around the construction and use of transition graphs of various kinds. To the best of my knowledge, transition graphs were first introducted in the NEWTON program of de Kleer (1977) and the FROB program of Forbus (1980). These programs addressed the problem of qualitative reasoning about the motion of point objects in various physical environments. The transition graphs they generated combined spatial continuity constraints with dynamic constraints.

Transition graphs of this kind were likewise the chief output of so-called “qualitative reasoning” systems (Kuipers, 1986; Forbus, 1985; de Kleer & Brown, 1985). These constructed transition graphs that characterize the behavior of dynamic systems whose state is a tuple of real-valued parameters. Again, the transition graphs in these systems combine dynamic constraints with continuity constraints; the first analysis to separate these was in section 4.8 of (Davis, 1990). The most relevant aspect of this work to the current chapter is the analysis of the topological issues, particularly the "transition ordering rules" of Williams (1985), and the "epsilon transition rule" of Kuipers (1986).

Configuration spaces were introduced as a method of characterizing the feasible behaviors of jointed robots in Lozano-Perez (1983). The problem of calculating the configuration space for interacting solid objects, given precise shape specifications, is known as the "motion planning" or "piano movers" problem. There is a large literature on the subject; see Part II of LaValle (2006) for an overview.

The RCC-8 relations between spatial regions were introduced in Randell and Cohn (1989). Figure 1 (with undirected arcs) was presented in Randell, Cui, and Cohn (1992) as showing the possible "topological" transitions between RCC relations—that is, those that are consistent with continuity constraints.

Galton (1993) presented the undirected transition graphs for RCC relations between two spatial fluents that are rigid but may interpenetrate. There are six of these, depending on the shapes of the two objects. Galton (1995) changed the undirected arcs representing transitions in previous studies to directed arcs; there is a directed arc from relation R to relation S if it is possible for R to hold at time *t* and S to hold at an open interval with lower bound *t* (In Galton's terminology, R "dominates" S).

Recent knowledge representation research on continuous spatial change has mostly followed either an *axiomatic* or a *semantic* approach. In the axiomatic approach, the objective is to characterize continuous change in terms of axioms over regions in space-time. The first attempt at this was in Muller (1998a). Muller used a first-order language over "histories" (Hayes, 1979)—that is, 4-dimensional regular regions in space-time—with three primitives: "C*xy*," meaning that histories *x* and *y* are connected; "$x<y$ " meaning that *x* strictly precedes *y* in time; and " $x \Diamond y$ " meaning that *x* and *y* overlap in time. Using this language, he proposed the following definition of "continuity":

$$\mathrm{CONTINU}w \triangleq$$
$$\mathrm{CON}_t w \wedge \forall_x \forall_u (\mathrm{TS}xw \wedge x \Diamond u \wedge \mathrm{P}uw) \Rightarrow \mathrm{C}xu.$$

Here "$\mathrm{CON}_t w$" means that the temporal projection of *w* is a connected time interval. "P*uw*" means that *u* is a subregion of *w*. "TS*xw*" means that *x* is a "time-slice" of *w*. All of these predicates can be defined in terms of the primitives (Muller, 1998a). It can be shown that Muller's axiomatic definition of continuity corresponded to continuity relative to the Hausdorff distance (Davis, 2001). In Muller (1998b) Muller observed that the above definition allows a 4-dimensional region that shrinks to a point at an instant and then expand from there to be considered continuous; he therefore proposed an alternative, stronger definition[1] of transitions.

Muller (1998b) also proposed an analysis of the feasible transitions between spatial RCC relations. For each RCC-8 relation R, he defines a spatial equivalent $\mathrm{R}_{sp}xy$, meaning that *x* and *y* are coextensive in time and that they are related spatially by R throughout their lifetimes. He was able to prove axiomatically the impossibility of a number of transitions for continuous regions.

However, Muller's approach does not allow the analysis of transitions that involve spatial relations that hold only for an instant. The limitation was addressed in Davis (2000), which gives first-order constructions that in effect define the cross-section C of a history at a given time, a point x in space-time, and the relations $\mathrm{x} \in$ Interior(*C*), and $\mathrm{x} \in$ Bd(*C*). The spatial RCC relations can then be defined in terms of points in the usual way, and the transition rules can be stated. However, these definitions are very complex, and it is unlikely that the transition rules can be *proven* from any plausible set of simple RCC axioms.

A more principled approach is taken in Hazarika and Cohn (2011). Here, the spatial relation between histories *x* and *y* at time *t* is defined in terms of the connectivity relations of $x \cap y$, $x \cup y$, $x - y$, and $y - x$ restricted to an interval ending at *t* and an interval beginning at *t*. Remarkably, using these definitions, they have been able to generate automatic proofs of the impossibility of each of the 45 transitions excluded in Figure 1 using the first-order theorem prover SPASS (Weidenbach, 2001).

In the semantic approach, continuity is characterized by defining a topology (generally a metric) over the space of spatial configurations. This approach was first applied to the analysis of transitions of RCC relations, in which a configuration is a pair of regular regions, in Galton (2000a) and Davis (2001), which independently arrived at very similar results. Galton considered five metrics over regular regions: the Hausdorff distance, the Hausdorff distance between the boundaries (which, however, was found to be inadequate), the dual Hausdorff distance, the area of the symmetric distance, and the Fréchet distance between the boundaries. Davis considered four: the Hausdorff distance, the dual Hausdorff distance, the area of the symmetric difference, and the optimal-homeomorphism metric.

Galton's monograph "Qualitative Spatial Change" (Galton, 2000a) is an extensive and rich study, combining representational and philosophical analysis. Chapters 7 and 8 deal with continuity.

There has also been work on continuous change in discrete models of space. These, of course, require substantially different definitions, both of topological relations and of continuity (Galton, 2000b, 2003)

Finally, the concept of qualitative homogeneous decompositions introduced here is modelled on the concept of cylindrical algebraic decompositions in computational algebra (Collins, 1975), in the sense that a QHD, like a CAD, is a partition of the space into cells, each of which is connected and uniform with respect to specified properties.

3. QUALITATIVE HOMOGENEOUS DECOMPOSITIONS

In this section, we define QHD's and study their properties in an abstract and general setting. We will return to specific issues of geometric change in section 4.

The general problem we address here is to characterize a path through a general topological space **T** in terms of its transitions through "qualitatively different" subsets of **T**. For example **T** might be the space of all pairs of regular regions in the plane divided into eight subsets corresponding to one of the eight RCC relations. For the purposes of this section, we will treat the partition of **T** into "qualitatively different" subsets as externally given; that is, we are given a partition of **T** and told that this constitutes a qualitative discrimination. Our analysis below consists in subdividing this initial partition into a finer partition that characterizes points in terms of the qualitative characteristics of the paths that pass through them.

Throughout this section **T** will be a topological space; in some cases, we will impose stronger requirements. We will use boldface lower case letters, such as **x** to denote points in **T**; boldface upper case letters, such as **P**, to denote subsets of **T**; and calligraphic letters such as $\mathcal{U}$ to denotes sets of subsets of **T**.

Definition 1 *A collection* $\mathcal{U}$ of subsets of **T** *is a* partition *of* **T**, *if for every point* $\mathbf{x} \in \mathbf{T}$ there is exactly one $\mathbf{U} \in \mathcal{U}$ such that $\mathbf{x} \in \mathbf{U}$. This set **U** *is called the "owner" of* **x** *in* $\mathcal{U}$, denoted "$\mathbf{O}(\mathbf{x}, \mathcal{U})$." The sets in $\mathcal{U}$ are called the cells of $\mathcal{U}$.

Definition 2 *Let* $\mathcal{U}$ and $\mathcal{V}$ be partitions of **T**. $\mathcal{V}$ is a (non-strict) refinement *of* $\mathcal{U}$, if for every $\mathbf{V} \in \mathcal{V}$ there exists $\mathbf{U} \in \mathcal{U}$ such that $\mathbf{V} \subset \mathbf{U}$.

Definition 3 *A* path *through* **T** *is a continuous function from the closed real interval [0,1] to* **T.** *A subset* **U** *of* **T** *is* path-connected *if, for every* $\mathbf{x}, \mathbf{y} \in \mathbf{U}$, there exists a path from **x** *through* **U** *to* **y**.

Note that the image of a path in **T** is a compact set in **T**.

Definition 4 *A partition* $\mathcal{U}$ of **T** *is* locally finite *if, for every point* $\mathbf{x} \in \mathbf{T}$ there exists a neighborhood **N** *of* x such that **N** *intersects only finitely many cells of* $\mathcal{U}$.

Definition 5 *Let* $\mathcal{U}$ be a locally finite partition of **T**. *The set of* neighbors *of* **x** *in* $\mathcal{U}$, denoted " $\mathbf{N}(\mathbf{x},\mathcal{U})$ " is defined as:

$$\mathbf{N}(\mathbf{x},\mathcal{U}) = \{\mathbf{U} \in \mathcal{U} \mid \mathbf{x} \in Cl_{\mathbf{T}}(\mathbf{U})\}$$

where $Cl_{\mathbf{T}}(\mathbf{U})$ is the topological closure of **U** with respect to **T**.

Note that the owner of **x** is always one of its neighbors. If **x** is in the interior of its owner, then its owner is the only neighbor. In particular, if **U** is open, then **U** is the only neighbor of any of the points in **U**. If **x** is on the boundary of its owner, then **x** has other neighbors in addition to its owner.

Definition 6 *Let* $\mathcal{U}$ be a locally finite partition of **T**. *A cell* $\mathbf{U} \in \mathcal{U}$ is uniform in its neighbors *if, for every two points* $\mathbf{x},\mathbf{y} \in \mathbf{U}$, $\mathbf{N}(\mathbf{x},\mathcal{U}) = \mathbf{N}(\mathbf{y},\mathcal{U})$.

Definition 7 *A path* π has a starting transition *from point* **x** *to set* **V** *if* $\pi(0) = \mathbf{x}$, and for all $t \in (0,1]$, $\pi(t) \in \mathbf{V}$. π has an ending transition *from set* **V** *to point* **x** *if* $\pi(1) = \mathbf{x}$, and for all $t \in [0,1)$, $\pi(t) \in \mathbf{V}$.

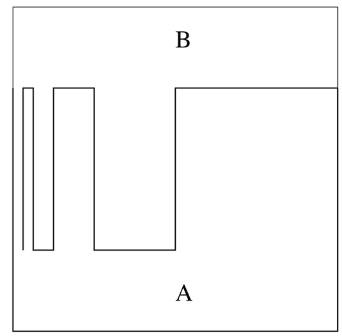
Definition 8 *Let* $\mathcal{U}$ be a locally finite partition of **T**. $\mathcal{U}$ allows simple transitions *if, for every point* $\mathbf{x} \in \mathbf{T}$ and for every neighbor **V** *of* **x** *in* $\mathcal{U}$, there exists a path π with a starting transition from **x** *to* **V**.

An example of a partition that is locally finite but does not allow simple transitions is as follows: Let **T** be the closed unit square. Let:

$$\mathbf{A} = ([0,1]\times[0,1/4]) \cup ([0,0]\times[0,3/4]) \cup \bigcup_{k=0}^{\infty}[1/4,3/4]\times[2^{-(2k+1)},2^{-2k}]$$

$$\mathbf{B} = \mathbf{T} - \mathbf{A}.$$

Figure 2. A locally finite partition that does not allow simple transitions



Thus **A** and **B** are alternating combs with infinitely many teeth (Figure 2). Note that **A** and **B** are both path-connected, **A** is closed regular, and **B** is open in **T**. The point $\langle 0,1/2 \rangle$ is in **A** and is in the closure of **B**, but there is no starting transition from that point to **B**.

Another, more outré, example: Let **T** be the unit square [0,1] $\times$ [0,1]. Define **A** and **B** as follows:

$$\mathbf{A} = \{\langle x,y \rangle \mid y = 0 \vee (0 < y < 1 \wedge x \text{ is rational}) \}$$

$$\mathbf{B} = \{\langle x,y \rangle \mid y = 1 \vee (0 < y < 1 \wedge x \text{ is irrational} \}$$

That is, **A** is the line at the bottom of the square plus all vertical lines with rational *x*-coordinates. **B** is the line at the top of the square plus all vertical lines with irrational *x*-coordinates. Then **A** and **B** are path-connected and uniform in their neighbors; every point has the neighbor set $\{\mathbf{A},\mathbf{B}\}$. However, there cannot exist a path that has a starting transition from any point in **A** to **B**, or from any point in **B** to **A**, except where the starting point is at the boundary of the square.

Definition 9 *A partition is* locally simple *if it is locally finite and allows simple transitions.*

Definition 10 *Let* $\mathcal{U}$ be a locally simple partition of **T**. $\mathcal{U}$ is a path-connected partition *of* **T** *if every cell* $\mathbf{U} \in \mathcal{U}$ is path-connected. $\mathcal{U}$ is a homogeneous partition *of* **T** *if every cell* $\mathbf{U} \in \mathcal{U}$ is path-connected and uniform in its neighbors.

Definition 11 *Let* $\mathcal{U}$ and $\mathcal{Q}$ be partitions of **T**. $\mathcal{Q}$ is the qualitative homogeneous decomposition (QHD) *of* $\mathcal{U}$ if the following are satisfied:

a. $\mathcal{Q}$ *is a homogeneous refinement of* $\mathcal{U}$.
b. *Every homogeneous refinement of* $\mathcal{U}$ *is a refinement of* $\mathcal{Q}$.

That is, $\mathcal{Q}$ is the coarsest homogeneous refinement of $\mathcal{U}$.

It is clear from the definition that any partition $\mathcal{U}$ can have at most one QHD (if there were two, each would be a refinement of the other); hence the phrase "*the* QHD of $\mathcal{U}$" is justified. A partition $\mathcal{U}$ may have no QHD; for example, there may exists a point **x** that is in the closure of infinitely many connected components of $\mathcal{U}$. However, theorem 1 states that, if $\mathcal{U}$ has any homogeneous refinement, then it has a QHD.

Theorem 1 *Let* $\mathcal{U}$ be a partition of **T**. *If there exists a homogeneous refinement of* $\mathcal{U}$, then there exists a QHD of $\mathcal{U}$.

Proof: See Appendix.

Theorem 2 gives a "constructive" definition for the QHD using transfinite induction.

Definition 12 *Let* $\mathcal{U}$ be a locally simple partition of **T**. *Define the function* $\Phi(\mathcal{U})$ to be the collection of all the path-connected components of $\mathcal{U}$. Define the equivalence relation $\mathbf{x} \sim_U \mathbf{y}$ as holding if **x** and **y** have the same owner and the same neighbors in $\mathcal{U}$; that is, $\mathbf{O}(\mathbf{x},\mathcal{U}) = \mathbf{O}(\mathbf{y},\mathcal{U})$ and $N(x,\mathcal{U}) = N(y,\mathcal{U})$. Define the function $\Psi(\mathcal{U})$ as the collection of equivalence classes of **T** *under the relation* $\sim_U$.

A number of immediate consequences may be noted. First, $\Phi(\mathcal{U})$ and $\Psi(\mathcal{U})$ are refinements of $\mathcal{U}$. Second, $\Phi(\mathcal{U})$ is the coarsest path-connected refinement of $\mathcal{U}$; it is the set of *path-connected components* of $\mathcal{U}$. Third, a locally simple partition $\mathcal{V}$ is homogeneous if it is a fixed point under Φ and Ψ; that is, $\Phi(\mathcal{V}) = \Psi(\mathcal{V}) = \mathcal{V}$.

Definition 13 *Let* $\mathcal{U}$ be a locally simple partition of **T**. *The* decompositional sequence corresponding to $\mathcal{U}$ is a sequence of refinements $\mathcal{U}_\sigma$ indexed by ordinals σ as follows:

- $\mathcal{U}_0 = \mathcal{U}$.
- *For each ordinal* σ, $\mathcal{U}_{\sigma+1} = \Psi(\Phi(\mathcal{U}_\sigma))$, where $\sigma + 1$ is the successor to σ
- *For each limit ordinal* σ, define the equivalence relation over **T**, $\mathbf{x} \sim_\sigma \mathbf{y}$ if, for all $i < \sigma$, $\mathbf{O}(\mathbf{x},\mathcal{U}_i) = \mathbf{O}(\mathbf{y},\mathcal{U}_i)$. Define $\mathcal{U}_\sigma$ to be the equivalence classes of **T** *under* $\sim_\sigma$.

Theorem 2 *Let* $\mathcal{U}$ be a locally simple partition of **T**. *Let* $\mathcal{U}_\sigma$ be the decompositional sequence of $\mathcal{U}$. Then:

- *The sequence reaches a fixed point. That is, there exists an ordinal* τ such that, for all $\sigma > \tau$, $\mathcal{U}_\sigma = \mathcal{U}_\tau$.
- *If* $\mathcal{U}_\tau$ is locally simple, then it is the QHD of $\mathcal{U}$.
- *If there exists a QHD of* $\mathcal{U}$, then it is $\mathcal{U}_\tau$.

Proof: See Appendix.

Tables 1 and 2 show simple examples of decompositional sequences, illustrated in Figures 3 and 4.

In both these examples, the final partition divides the points in **T** according to the local topological structure of the starting partition; that is, in the final QHD, any two points in the same cell have neighborhoods that are homeomorphic in terms of the original labels. That is not always the case: for instance, if **A** is the line from $\langle -1, 0\rangle$ to $\langle 1, 0\rangle$, and **B** is the complement of **A**, then {**A**,**B**} is a homogeneous partition, even though the neighborhoods of the end points of the line are not homemorphic to the neighborhoods of the interior points of the line. The analysis in section 4.1 of the three-dimensional jigsaw puzzle will give another example of this.

Table 1. Decompositional sequence: example 1

Let **T** be the plane, and let $\mathcal{U}$ consist of two cells:

A: The union of the solid disk of radius 1 centered at $\langle 1,0 \rangle$, the annulus with inner radius 1/2 and outer radius 1 centered at $\langle -1,0 \rangle$, and the solid disk of radius 1 centered at $\langle 0,2 \rangle$.
B: The complement of **A**.

The decompositional sequence proceeds as follows:

$\mathcal{U}_0 = \mathcal{U}$.

$\Phi(\mathcal{U}_0) = \{\mathbf{A1}, \mathbf{A2}, \mathbf{B1}, \mathbf{B2}\}$ where
$\mathbf{A1}$ is the disk centered at $\langle 0,2 \rangle$.
$\mathbf{A2} = \mathbf{A} - \mathbf{A1}$ is the union of the other disk and the annulus;
$\mathbf{B1}$ is the open disk of radius 1/2 centered at $\langle -1,0 \rangle$
$\mathbf{B2} = \mathbf{B} - \mathbf{B1}$ is the exterior of $\mathbf{A}$.
$\mathcal{U}_1 = \Psi(\Phi(\mathcal{U}_0)) = \{\mathbf{A1a}, \mathbf{A1b}, \mathbf{A2a}, \mathbf{A2b}, \mathbf{A2c}, \mathbf{B1}, \mathbf{B2}\}$ where
$\mathbf{A1a}$ is the boundary of $\mathbf{A1}$. Neighbor set: { $\mathbf{A1, B2}$ }.
$\mathbf{A1b}$ is the interior of $\mathbf{A1}$. Neighbor set: { $\mathbf{A1}$ }.
$\mathbf{A2a}$ is the part of boundary of $\mathbf{A2}$ bordering $\mathbf{B2}$ (the Figure 8).
Neighbor set: { $\mathbf{A2, B2}$ }.
$\mathbf{A2b}$ is the interior of $\mathbf{A2}$. Neighbor set: $\mathbf{A2}$.
$\mathbf{A2c}$ is the part of the boundary of $\mathbf{A2}$ bordering $\mathbf{B1}$ (the inner circle).
Neighbor set: { $\mathbf{A2, B1}$ }.

$\Phi(\mathcal{U}_1) = \{\mathbf{A1a}, \mathbf{A1b}, \mathbf{A2a}, \mathbf{A2b1}, \mathbf{A2b2}, \mathbf{A2c}, \mathbf{B1}, \mathbf{B2}\}$ where
$\mathbf{A2b1}$ is the interior of the right-hand disk.
$\mathbf{A2b2}$ is the interior of the annulus.
$\mathcal{U}_2 = \Psi(\Phi(\mathcal{U}_1)) = \{\mathbf{A1a}, \mathbf{A1b}, \mathbf{A2a1}, \mathbf{A2a2}, \mathbf{A2a3}, \mathbf{A2b1}, \mathbf{A2b2}, \mathbf{A2c}, \mathbf{B1}, \mathbf{B2}\}$.
$\mathbf{A2a1}$ is the point $\langle 0,0 \rangle$. Neighbor set: $\{\mathbf{A2a}, \mathbf{A2b1}, \mathbf{A2b2}, \mathbf{B2}\}$.
$\mathbf{A2a2}$ is the right-hand circle except for $\langle 0,0 \rangle$. Neighbor set: $\{\mathbf{A2a}, \mathbf{A2b1}, \mathbf{B2}\}$.
$\mathbf{A2a3}$ is the left-hand circle except for $\langle 0,0 \rangle$. Neighbor set: $\{\mathbf{A2a}, \mathbf{A2b2}, \mathbf{B2}\}$.

$\Phi(\mathcal{U}_2) = \mathcal{U}_2$.
$\mathcal{U}_3 = \Psi(\Phi(\mathcal{U}_2)) = \mathcal{U}_2$, so a fixed point has been reached.

3.1. Transition Graphs

The structure of a locally simple partition can be expressed as a graph. It should be noted that these "graphs" are not quite standard, in that they may have infinitely many vertices.

Definition 14 *Let* $\mathcal{U}$ be a partition of **T**. *The corresponding* transition graph *G is defined as follows:*

- *The vertices of G* are the cells in $\mathcal{U}$.
- *For any* $\mathbf{U}, \mathbf{V} \in \mathcal{U}$, there is an arc from **U** *to* **V** *if* $\mathbf{V} \neq \mathbf{U}$ and $\mathbf{U} \cap Cl_T(\mathbf{V})$ is non-empty. (That is, **V** *is a neighbor of some point* $\mathbf{x} \in \mathbf{U}$.)

In the terminology of Galton (1995), there is an arc from **U** to **V** if **U** dominates **V**.

In general, a vertex in a transition graph may have infinite in-degree and infinite out-degree.

Table 2. Decompositional sequence: example 2

Let **T** be the plane, and let $\mathcal{U}$ consist of two cells:

A: The upper half-disk, open on the bottom. That is, $\mathbf{A} = \{\langle x, y\rangle \mid x^2 + y^2 \leq 1 \wedge y > 0\}$
B: The complement of **A**.

The decompositional sequence proceeds as follows:

$\mathcal{U}_0 = \mathcal{U}$.

$\Phi(\mathcal{U}_0) = \mathcal{U}_0$, since **A** and **B** are each path-connected.
$\mathcal{U}_1 = \Psi(\Phi(\mathcal{U}_0)) = \{\mathbf{A1}, \mathbf{A2}, \mathbf{B1}, \mathbf{B2}\}$ where
$\mathbf{A1}$ is the interior of the half-disk. Neighbor set $\{\mathbf{A}\}$
$\mathbf{A2}$ is the semi-circle $\mathbf{A} = \{\langle x, y\rangle \mid x^2 + y^2 = 1 \wedge y > 0\}$. Neighbor set $\{\mathbf{A}, \mathbf{B}\}$.
$\mathbf{B1}$ is the line from $\langle -1, 0\rangle$ to $\langle 1, 0\rangle$. Neighbor set $\{\mathbf{A}, \mathbf{B}\}$.
$\mathbf{B2} = \mathbf{B} - \mathbf{B1}$. Neighbor set $\{\mathbf{B}\}$.

$\Phi(\mathcal{U}_1) = \mathcal{U}_1$.
$\mathcal{U}_2 = \Psi(\Phi(\mathcal{U}_1)) = \{\mathbf{A1}, \mathbf{A2}, \mathbf{B1a}, \mathbf{B1b}, \mathbf{B2}\}$, where
$\mathbf{B1a}$ is the two endpoints of the line, $\{\langle -1, 0\rangle, \langle 1, 0\rangle\}$.
Neighbor set $\{\mathbf{A1}, \mathbf{A2}, \mathbf{B1}, \mathbf{B2}\}$
$\mathbf{B1b} = \mathbf{B1} - \mathbf{B1a}$ is the rest of the line. Neighbor set $\{\mathbf{A1}, \mathbf{B1}, \mathbf{B2}\}$.

$\Phi(\mathcal{U}_2) = \{\mathbf{A1}, \mathbf{A2}, \mathbf{B1a1}, \mathbf{B1a2}, \mathbf{B1b}, \mathbf{B2}\}$ where
$\mathbf{B1a1} = \{\langle -1, 0\rangle\}$, the left-hand end point.
$\mathbf{B1a2} = \{\langle 1, 0\rangle\}$, the right-hand end point.
$\mathcal{U}_3 = \Psi(\Phi(\mathcal{U}_2)) = \Phi(\mathcal{U}_2)$.

$\Phi(\mathcal{U}_3) = \mathcal{U}_3$.
$\mathcal{U}_4 = \Psi(\Phi(\mathcal{U}_3)) = \mathcal{U}_3$, so a fixed point has been reached.

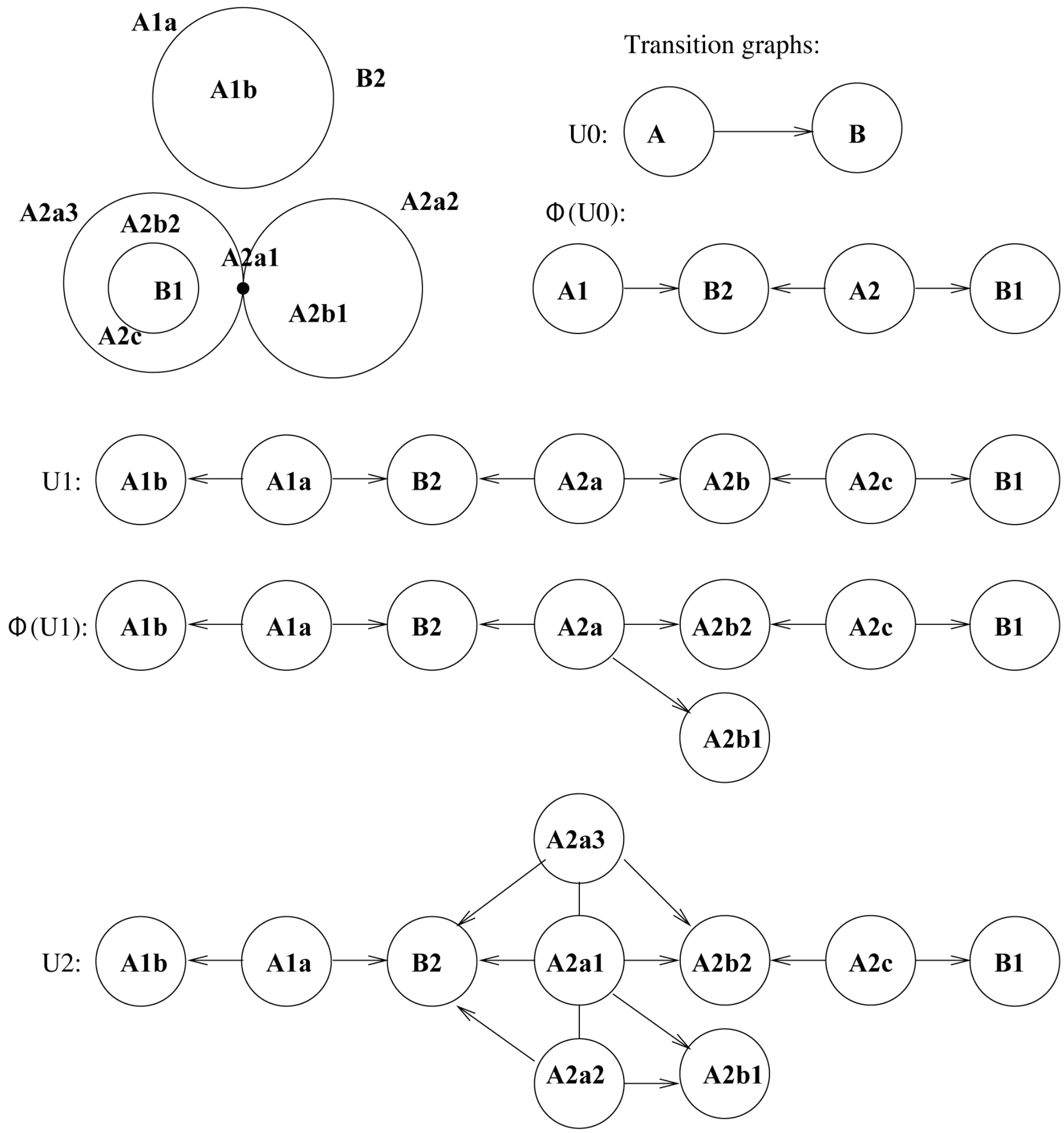
If the partition is homogeneous, then since all points in the cell have the same neighbors and since a single point can have only finitely many neighbors, the out-degree of any vertex is finite, though the in-degree may still be infinite.

The direction of an arc in a transition graph does not indicate the allowable direction of a transition; a transition may occur either forward or backward along an arc. Rather, the direction of the arc indicates the *topology* of a transition. If there is an arc from cell **U** to **V**, then a path π carries out a forward transition along that arc at time t if $\pi(t) \in \mathbf{U}$, and $\pi(t') \in \mathbf{V}$ for $t' \in (t, t1)$ for some $t1 > t$. The path π carries out a backward transition along that arc at time t if $\pi(t') \in \mathbf{V}$ for $t' \in (t1, t)$ for some $t1 < t$ and $\pi(t) \in \mathbf{U}$.

Definition 15 *Let* $\mathcal{U}$ be a partition of **T**. *The* Qualitative Homogeneous Decomposition graph (QHD graph) *of* $\mathcal{U}$ is the transition graph corresponding to the QHD of $\mathcal{U}$.

There can exist partitions in which there are arcs in both directions between two cells. For instance, in the initial partition $\mathcal{U}$ in the example in Table 2, there is an arc from **A** to **B** and an arc from **B** to **A**. I do not know whether this can happen with a homogeneous partition over $\mathbb{R}^k$; it would certainly have to be highly pathological. It can happen with some kinds of partitions in

Figure 3. Decompositional sequence: example 1

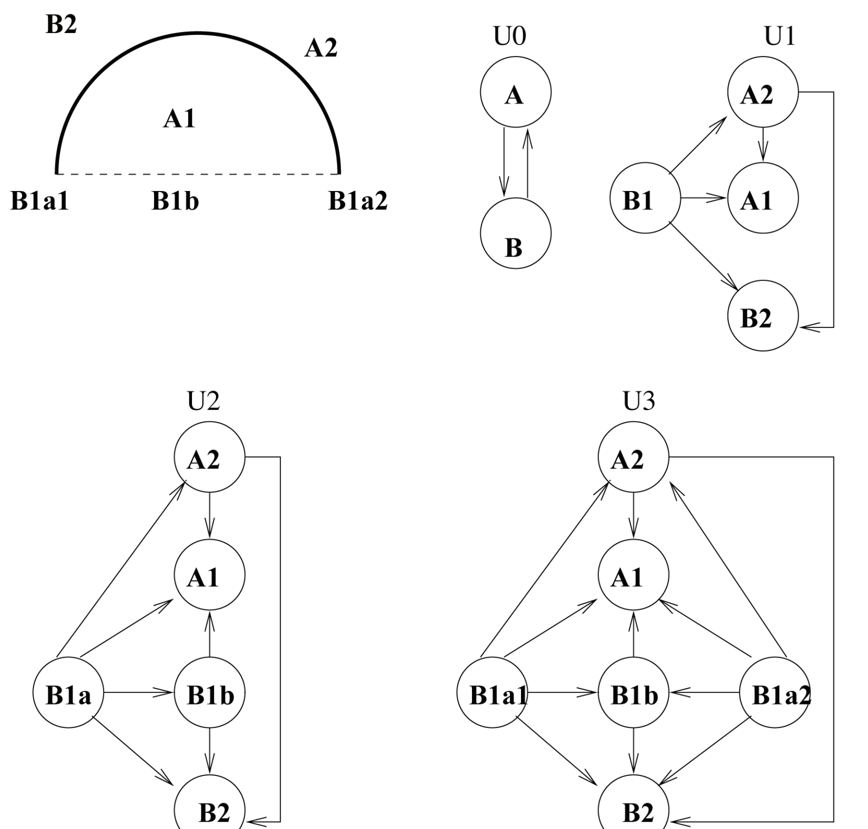


some topological spaces of regions over $\mathbb{R}^k$; for instance, in the space of open regular regions in the plane, topologized by the area of the symmetric regions, it is possible to have closed transitions in either direction between DC and EC, or between TPP and NTTP (Davis, 2001, Figure 9).

In such cases, we must distinguish between going from **A** to **B** forward along the arc $\langle \mathbf{A}, \mathbf{B} \rangle$ as opposed to going from **A** to **B** backward against the arc $\langle \mathbf{B}, \mathbf{A} \rangle$. In the first case, the path will be in **A** at the moment of transition, and in the second case it will be in **B**. Therefore, our definition of a path through the transition graph, which we will call a "gpath," is a little different from the usual:

Definition 16 *Let G be a transition graph. A* gpath *is an alternating sequence of vertices and edges in G* $\langle \mathbf{V}_1, A_1, \mathbf{V}_2, A_2, \ldots A_{k-1}, \mathbf{V}_k \rangle$, starting and ending with a vertex, such that A_i is either the

Figure 4. Decompositional sequence: example 2



arc $\langle \mathbf{V}_i, \mathbf{V}_{i+1} \rangle$ or the arc $\langle \mathbf{V}_{i+1}, \mathbf{V}_i \rangle$. The sequence of a single vertex $\langle \mathbf{V}_1 \rangle$ is also considered a gpath.

3.2. The Expressivity of QHD Graph

In this section, we study the relation between the paths through a partitioned space and the gpaths through the corresponding transition graph.

It will be convenient to define an intermediate structure, called an "Interval Label Sequence" (ILS). We will use π, ϕ, and ψ, as variables over paths, α for interval label sequences, and β for gpaths. We consider the vertices of the graph to be literally the cells of the partition; thus, we will use boldface letters for vertices of the graph.

We begin with a few definitions that describe how a path π through $\mathbf{T}$ is characterized in terms of the sequence of qualitative states it passes through.

A path through a partitioned space divides the unit time interval into subintervals; in each subinterval, the path remains in one cell. We are interested in the topology of these intervals. We consider two ways of characterizing the topology of a real interval, called "interval label sets." The Z4 label set includes four labels: open (`O'), closed (`C'), closed on the left (`L'), and closed on the right (`R'). The Z5 label set includes five labels; it divides closed intervals into closed instantaneous (`CI') and closed extended (`CE'), together with `O,' `L,' and `R.' The difference reflects the difference between characterizing the shape of successive *transitions,* in which case `CI' and `CE' are the same, versus characterizing the shape of the *intervals,* in which case they are different. As we shall see, the difference also corresponds to the difference between using a path-connected partition and a homogeneous partition.

Definition 17 *Let Z be a interval label set; that is, either Z4 or Z5 as described above. The* shape *of an interval I in label set Z is denoted "shape(* I, Z *)."*

For any bounded real interval *I*, let *l*(*I*) and $u(I)$ be respectively the lower and upper bounds of *I*.

Definition 18 *A finite interval partition of the interval [0,1] is a finite partition* $\langle I_1 \ldots I_k \rangle$ of [0,1] such that for $i = 1 \ldots k-1$, $u(I_i) = l(I_{i+1})$.

Definition 19 *Let* $\mathcal{U}$ be a partition of **T** *and let* π be a path through **T**. *The finite interval partition* $\langle I_1 \ldots I_k \rangle$ of [0,1] is induced by π with respect to $\mathcal{U}$ if π occupies the same set in $\mathcal{U}$ throughout each subinterval I_i and moves from one set to another in each transition from I_i to I_{i+1}. Formally,

- *For* $i = 1 \ldots k$, if $t_1, t_2 \in I_i$ then $\mathbf{O}(\pi(t_1), \mathcal{U}) = \mathbf{O}(\pi(t_2), \mathcal{U})$; and
- *For* $i = 1 \ldots k-1$, if $t_1 \in I_i, t_2 \in I_{i+1}$ then $\mathbf{O}(\pi(t_1), \mathcal{U}) \neq \mathbf{O}(\pi(t_2), \mathcal{U})$.

If π induces a finite interval partition, then it is said to be *finitary*. Not all paths are finitary; a path that moves infinitely often between sets in $\mathcal{U}$ does not induce a finite interval partition. We will limit our discussion to finitary paths.

Definition 20 *Let* $\mathcal{U}$ be a locally simple partition of **T** *and let* Z be an interval label set. An interval label pair *is a pair of a cell of* $\mathcal{U}$ and a label in Z. An Interval Label Sequence (ILS*) is a finite sequence* $\langle\langle U_1, z_1 \rangle \ldots \langle U_k, z_k \rangle\rangle$ of interval label pairs.

Definition 21 *Let* $\mathcal{U}$ be a locally simple partition of **T**, *let Z be an interval label set, and let* π be a finitary path. Let $\langle I_1 \ldots I_k \rangle$ be the interval partition induced by π. The interval trace *of* π through $\mathcal{U}$, denoted $\Gamma_{U,Z}(\pi)$, is the interval label sequence:

$$\begin{aligned}&\langle\langle \mathbf{O}(\pi(I_1), \mathcal{U}), shape(I_1, Z) \rangle, \\ &\quad \langle \mathbf{O}(\pi(I_2), \mathcal{U}), shape(I_2, Z) \rangle, \\ &\quad \ldots \langle \mathbf{O}(\pi(I_k), \mathcal{U}), shape(I_k, Z) \rangle\rangle\end{aligned} .$$

We have slightly abused notation in writing $\mathbf{O}(\pi(I_i), \mathcal{U})$ to mean the value of $\mathbf{O}(\pi(t), \mathcal{U})$ for all $t \in I_i$.

Definition 22 *The* closed labels *are `C,' `CI,' and `CE.' The* left-closed *labels are the closed labels and `L.' The* right-closed *labels are the closed labels and `R.' The* left-open *labels are `R' and `O.' The* right-open *labels are `L' and `O.'*

Definition 23 *An interval label sequence* $\langle\langle \mathbf{U}_1, z_1 \rangle \ldots \langle \mathbf{U}_k, z_k \rangle\rangle$ is coherent *if either*

a. $k = 1$ and $z_1 =$ `C' or `CE'; or
b. $k > 1$ and all the following hold:
   b.1 z_1 is left-closed.
   b.2 z_k is right-closed.
   b.3 for $i = 1 \ldots k-1$, either:
   b.3.a z_i is right-closed, z_{i+1} is left-open, and $\mathbf{U}_{i+1}$ is a neighbor of $\mathbf{U}_i$; or
   b.3.b z_i is right-open, z_{i+1} is left-closed, and $\mathbf{U}_i$ is a neighbor of $\mathbf{U}_{i+1}$.

Theorem 3 *The interval trace of any finitary path* π is a coherent ILS.

Proof: Straightforward.

Definition 24 *Let* $\mathcal{U}$ be a locally simple partition of **T**, *let G be the transition graph for* $\mathcal{U}$, and let Z be a label set. Let $\alpha = \langle\langle \mathbf{U}_1, z_1 \rangle \ldots \langle \mathbf{U}_k, z_k \rangle\rangle$ be a coherent ILS. The gpath through *G* corresponding to α, denoted $\Delta(\alpha)$, is the gpath $\beta = \langle \mathbf{U}_1, A_1, \ldots A_{k-1}, \mathbf{U}_k \rangle$ where $A_i = \langle \mathbf{U}_i, \mathbf{U}_{i+1} \rangle$ if z_i is right-closed and $A_i = \langle \mathbf{U}_{i+1}, \mathbf{U}_i \rangle$ if z_i is right-open.

Example: Consider the path π shown in Figure 5, using the partition of Figure 4. The corresponding gpath is:

$\langle$**B2**, $\langle$**B1a1**, **B2**$\rangle$, **B1a1**, $\langle$**B1a1**, **A1**$\rangle$, **A1**, $\langle$**B1b**, **A1**$\rangle$, **B1b**, $\langle$**B1b**, **A1**$\rangle$, **A1**, $\langle$**A2**, **A1**$\rangle$, **A2**, $\langle$**A2**, **B2**$\rangle$, **B2**$\rangle$.

The corresponding ILS is:

$\langle\langle \mathbf{B2}, `L'\rangle, \langle \mathbf{B1a1}, `CI'\rangle, \langle \mathbf{A1}, `O'\rangle, \langle \mathbf{B1b}, `CE'\rangle, \langle \mathbf{A1}, `O'\rangle, \langle \mathbf{A2}, `CI'\rangle, \langle \mathbf{B2}, `R'\rangle\rangle$

Theorem 4 *Under the assumptions of definition 24, if* α is a coherent ILS, then $\Delta(\alpha)$ is a gpath. Conversely, if β is a gpath then for both Z4 and Z5, there exists a coherent ILS α such that $\Delta(\alpha) = \beta$ (For Z4, α is unique. In Z5, it may not be; if arc A_{j-1} is traversed backward, and arc A_j is traversed forward, then vertex E_j may be labelled either `CI' or `CE.').

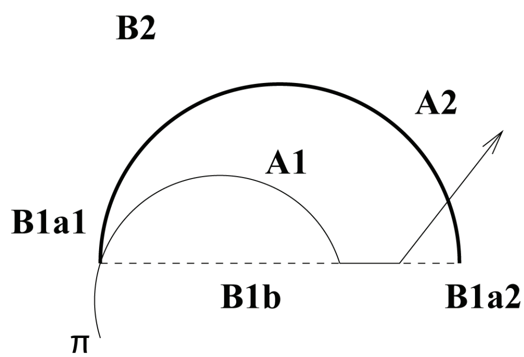
Proof: Immediate from the definitions.

Definition 25 *For any path* π through **T**, *the* start *of* π, $S(\pi) = \pi(0)$ and the end *of* π, $E(\pi) = \pi(1)$. For any interval label sequence $\alpha = \langle\langle \mathbf{V}_1, s_1\rangle \ldots \langle \mathbf{V}_k, s_k\rangle\rangle$, the start of α, $S(\alpha) = \mathbf{V}_1$ and the end of α, $E(\alpha) = \mathbf{V}_k$. For any gpath $\alpha = \langle \mathbf{V}_1, A_1, \ldots A_{k-1}, \mathbf{V}_k\rangle$, the start of α, $S(\alpha) = \mathbf{V}_1$ and the end of α, $E(\alpha) = \mathbf{V}_k$.

We now show the converse of theorems 3 and 4: Any coherent ILS or coherent gpath is the trace of a finitary path. Moreover, one can choose the starting and ending points of the path to be any points in the starting and ending cell of the ILS/gpath.

Theorem 5 *Let* $\mathcal{U}$ be a path-connected partition of **T**. *Let* α be a coherent Z4 ILS for $\mathcal{U}$. Let **x** *be a point in* $S(\alpha)$ and let **y** *be a point in* $E(\alpha)$. Then there exists a finitary path π through **T** *suchthat* $S(\pi) = \mathbf{x}$, $E(\pi) = \mathbf{y}$, and $\Gamma_{U,Z4}(\pi) = \alpha$.

Figure 5. Trace of a path



Proof: If $\alpha = \langle\langle \mathbf{U}, `C'\rangle\rangle$, then, since **U** is path-connected, let π be a path through **U** connecting **x** and **y**.

Otherwise, let $\alpha = \langle\langle \mathbf{U}_1, z_1\rangle \ldots \langle \mathbf{U}_k, z_k\rangle\rangle$. For $i = 2 \ldots k-1$, let $\mathbf{a}_i$ be a point in $\mathbf{U}_i$; let $\mathbf{a}_1 = \mathbf{x}$ and let $\mathbf{a}_k = \mathbf{y}$. For $i = 1 \ldots k-1$ if z_i is right-closed, let $\mathbf{b}_i$ be a point in $\mathbf{U}_i \cap \mathrm{Cl}(\mathbf{U}_{i+1})$. Since $\mathbf{a}_i$ and $\mathbf{b}_i$ are in $\mathbf{U}_i$, which is path-connected, there is a path π_i from $\mathbf{a}_i$ to $\mathbf{b}_i$. Since $\mathcal{U}$ is locally simple, there is a path ϕ_i that has a starting transition from $\mathbf{b}_i$ into $\mathbf{U}_{i+1}$. Since $\mathbf{U}_{i+1}$ is path-connected, there is a path ψ_i from $\phi_i(1)$ to $\mathbf{a}_{i+1}$. If z_i is right-open, let $\mathbf{b}_i$ be a point in $\mathrm{Cl}(\mathbf{U}_i) \cap \mathbf{U}_{i+1}$. Since $\mathbf{a}_{i+1}$ and $\mathbf{b}_i$ are in $\mathbf{U}_i$, which is path-connected, there is a path ψ_i from $\mathbf{b}_i$ to $\mathbf{a}_{i+1}$. Since $\mathcal{U}$ is locally simple, there is a path ϕ_i that has an ending transition from $\mathbf{U}_i$ to $\mathbf{b}_i$. Since $\mathbf{U}_i$ is path-connected, there is a path π_i from $\mathbf{a}_i$ to $\phi_i(0)$. Then splicing all these together, the path $\pi_1 \mid \phi_1 \mid \psi_1 \mid \pi_2 \mid \phi_2 \mid \psi_2 \mid \ldots \pi_k \mid \phi_k \mid \psi_k$ is a path satisfying the conclusion of the lemma.

Theorem 6 *Let* $\mathcal{U}$ be a homogeneous partition of **T**. *Let* α be a coherent Z5 ILS through $\mathcal{U}$, let **x** be a point in $S(\alpha)$ and let **y** be a point in $E(\alpha)$. Then there exists a finitary path π through **T** *such that* $S(\pi) = \mathbf{x}$, $E(\pi) = \mathbf{y}$, and $\Gamma_{U,Z5}(\pi) = \alpha$.

Proof: The proof is the same as that of the previous theorem, with two changes. First, since cells are uniform in their neighbors, the point $\mathbf{b}_i$

can be identified with $\mathbf{a}_i$ if z_i is right-closed, and with $\mathbf{a}_{i+1}$ if z_i is right-open, so the paths from $\mathbf{a}_i$ to $\mathbf{b_i}$ and from $\mathbf{b_i}$ to $\mathbf{a}_{i+1}$ can be omitted. Second, for any $\mathbf{a}_i$ let ψ_{i-1} be the path leading into $\mathbf{a}_i$ and let π_i be the path leading out of $\mathbf{a}_i$, as constructed in the proof of the previous lemma. If z_i is `CI,' then splice $\pi_i$ directly onto $\psi_i$. If $z_i$ is `CE,' then construct a path ϕ that remains at the point $\mathbf{a}_i$ for all of [0,1], and splice this between ψ_i and π_i. ▪

Corollary 7 *Let* $\mathcal{U}$ be a path-connected partition of **T**. *Let* Z be an interval label set. Let β be a gpath through the transition graph of $\mathcal{U}$. Let **x** *be a point in* $S(\beta)$ and let **y** be a point in $E(\beta)$. Then there exists a finitary path π through T such that $S(\pi) = \mathbf{x}$, $E(\pi) = \mathbf{y}$, and $\Delta(\Gamma_{U,Z}(\pi)) = \beta$.

Proof: Immediate from Theorems 4 and 5.

Note that corollary 7 holds even if $Z = Z5$ and $\mathcal{U}$ is not homogeneous, because the translation Δ from interval label sequences to gpaths obliterates the distinction between Z4 and Z5 labels.

We next show that the condition in theorem 5 that $\mathcal{U}$ is path-connected and the condition in theorem 6 are necessary conditions; the conclusion holds only if the conditions are satisfied:

Theorem 8 *Let* $\mathcal{U}$ be a locally simple partition of **T**. *Suppose it is true that, for every coherent Z4 ILS* α through $\mathcal{U}$, and for any points **x** in $S(\alpha)$, **y** in $E(\alpha)$, there exists a path π such that $S(\pi) = \mathbf{x}$, $E(\pi) = \mathbf{y}$, and $\Gamma_{U,Z4}(\pi) = \alpha$. Then $\mathcal{U}$ is path-connected.

Proof of the contrapositive. Let **U** be a cell in $\mathcal{U}$ that is not path-connected, and let **x** and **y** be points in different path-connected components of **U**. Then the interval label sequence $\langle \mathbf{U}, C \rangle$ is coherent, but there is no path from **x** to **y** through **U**.

Theorem 9 *Let* $\mathcal{U}$ be a locally simple partition of **T**. *Suppose it is true that for every coherent Z5 interval label sequence* α through $\mathcal{U}$, and for any points **x** in $S(\alpha)$, **y** in $E(\alpha)$, there exists a path π such that $S(\pi) = \mathbf{x}$, $E(\pi) = \mathbf{y}$, and $\Gamma_{U,Z5}(\pi) = \alpha$. Then $\mathcal{U}$ is homogeneous.

Proof of the contrapositive. Suppose that $\mathcal{U}$ is not homogeneous. Then there exists a cell **U** in $\mathcal{U}$ which is either not path connected or not uniform in its neighbors. If it is not path-connected, then the proof proceeds as in theorem 8. If it is not uniform in its neighbors, then there exists $\mathbf{x} \in \mathbf{U}$ and a neighbor **V** of **U** such that **V** is not a neighbor of **x**. Then the Z5 interval label sequence $\langle\langle \mathbf{U}, \text{`CI}' \rangle, \langle \mathbf{V}, \text{`R}' \rangle\rangle$ is coherent, but there is no path starting in **x** with that trace.

A Z5 interval label sequence fully characterizes the topology of an interval partition:

Definition 26 *Two interval partitions* $\langle I_1 \ldots I_k \rangle$ and $\langle J_1 \ldots J_k \rangle$ are homeomorphic *if and only if there is a direction-preserving automorphism* $\mathcal{H}(t)$ from [0,1] to [0,1] such that $\mathcal{H}(I_i) = J_i$.

Theorem 10 $\langle I_1 \ldots I_k \rangle$ and $\langle J_1 \ldots J_k \rangle$ are homeomorphic if shape($I_i, Z5$) = shape($J_i, Z5$) for $i = 1 \ldots k$.

Proof: It is immediately clear topologically that if $J_i = \mathcal{H}(I_i)$ then shape($I_i, Z5$) = shape($J_i, Z5$). To prove the converse, we may construct $\mathcal{H}(t)$ as follows: For any t, let I_i be the interval containing t. If I_i is an instantaneous interval, define $\mathcal{H}(t) = l(J_i)$. Otherwise, let:

$$\mathcal{H}(t) = l(J_i) + \frac{(t - l(I_i)) * (u(J_i) - l(J_i))}{u(I_i) - l(I_i)}$$

It is easily shown that $\mathcal{H}$ satisfies the conditions of the theorem.

Definition 27 *Paths* π and ϕ are $\mathcal{U}$-homeomorphic *if there exists a direction-preserving automorphism* $\mathcal{H}(t)$ from [0,1] to [0,1] such that, for all t, $\mathbf{O}(\pi(t), \mathcal{U}) = \mathbf{O}(\phi(\mathcal{H}(t)), \mathcal{U})$.

Theorem 11 *Paths* π and ϕ are $\mathcal{U}$-homeomorphic if and only if $\Gamma_{U,Z5}(\pi) = \Gamma_{U,Z5}(\phi)$.

Proof: The implication left to right is immediate from theorem 10. For the implication right to left, let $\langle I_1 \ldots I_k \rangle$ and $\langle J_1 \ldots J_k \rangle$ be the interval partitions induced by π and ϕ respectively, and let $\mathcal{H}$ be the function defined in the proof of theorem 10. It is immediate that this satisfies the conditions.

We now show that Γ preserves the operation of splicing two paths, suitably defined.

Definition 28 *Let* π and ϕ be paths such that $E(\pi) = S(\phi)$. The simple splice *of* π and ϕ, denoted $\pi \mid \phi$, is the path ψ such that:

$$\psi(t) = \begin{cases} \pi(2t) & for\ 0 \leq t \leq 1/2; \\ \phi(2t-1) & for\ 1/2 \leq t \leq 1. \end{cases}$$

Definition 29 *Let* z_1 and z_2 be labels from the same interval label set, where z_1 is closed-right, and z_2 is closed-left. The splice *of* z_1 and z_2 denoted $z_1 \mid z_2$ is defined as shown in Table 3.

Definition 30 *Let* $\alpha_1 = \langle \langle \mathbf{V}_1, z_1 \rangle \ldots \langle \mathbf{V}_k, z_k \rangle \rangle$ and $\alpha_2 = \langle \langle \mathbf{W}_1, y_1 \rangle \ldots \langle \mathbf{W}_m, y_m \rangle \rangle$ be two coherent interval label sequences from the same label set. If $\mathbf{V}_k = \mathbf{W}_1$ then the splice *of* α_1 and α_2,

$$\alpha_1 \mid \alpha_2 = \langle \langle \mathbf{V}_1, z_1 \rangle \ldots \langle \mathbf{V}_{k-1}, z_{k-1} \rangle \ldots \langle \mathbf{V}_k = \mathbf{W}_1, z_k \mid y_1 \rangle, \langle \mathbf{W}_2, y_2 \rangle \ldots \langle \mathbf{W}_m, y_m \rangle \rangle.$$

Note that since α_1 and α_2 are coherent, it follows that $z_k \mid y_1$ is defined.

Theorem 12 *If* π and ϕ are paths such that $E(\pi) = S(\phi)$ then $\Gamma_{U,Z}(\pi \mid \phi) = \Gamma_{U,Z}(\pi) \mid \Gamma_{U,Z}(\phi)$. If α_1 and α_2 are interval label sequences such that $E(\alpha_1) = S(\alpha_2)$ then $\Delta(\alpha_1 \mid \alpha_2) = \Delta(\alpha_1) \mid \Delta(\alpha_2)$.

Proof: Immediate from the definitions.

3.3. Metalogical Theorems

Using the above semantic theorems, we can now prove metalogical results, showing that the decision problem for certain first-order languages over the domains of finitary paths can be reduced to a decision problem for corresponding languages over gpaths and ILS's.

There are a couple of issues to address at the outset. First, one has to be careful here not to make the language of paths too expressive. In particular, if the language supplies any way to express the relation that a point lies in the *middle* of a path, then the jig is up; there is no way to achieve this kind of reduction. The problem is that, in a one-dimensional cell **U**, like **B1b** of Table 2, it is a fact that there exist three distinct points $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{U}$ such that any path from **x** to **y** that remains in **U** must go through **z**. In a language of points and paths with a predicate "On(x,p)," meaning point x lies on path p, this statement can be expressed in the formula

$\exists_{x,y,z}$ x≠z≠y ⋀ ∀p [On(x,p) ⋀On(y,p) ⋀ [∀w On(w,p) ⇒ In(w,U)]] ⇒ On(z,p).

This kind of constraint is quite hopeless to capture in transition graphs or any reasonable extension of them. Since the relation "On(x,p)" can be defined in terms of the splice function, one likewise has to exclude the splice function. However, one needs something like splice so that there is some way to build up extended gpaths. The way out is that we will use relations "Z4Splice(π, ϕ, ψ)" and "Z5Splice(π, ϕ, ψ)," which hold if $\Gamma_{U,Z}(\psi) = \Gamma_{U,Z}(\pi) \mid \Gamma_{U,Z}(\phi)$ for $Z = Z4$ or $Z5$, respectively.

Second, even with this limited language, there is an inescapable logical distinction between a cell **U** with a single point, such as **A2a1** of Table 1, which satisfies the formula "$\forall_{x,y}$ In(x,u) ∧ In(y,u) ⇒ x=y" and a cell with infinitely many points, which does not. (In a Hausdorff topology,[2] any path-connected cell must be one or the other.) Therefore, the cells in a transition graph must be labelled as either "singleton cells" or "infinite cells."

We can now proceed to our formal construction. First, more typographical conventions. We

Table 3. Same interval label set

$z1$ / $z2$	`C'	`CI'	`CE'	`L'
`C'	`C'	---	---	`L'
`CI'	---	`CI'	`CE'	`L'
`CE'	---	`CE'	`CE'	`L'
`R'	`R'	`R'	`R'	`O'

will use symbols in block capitals for relations in the domains. We will use typewriter font, such as Splice(a,b,c) for symbols in the formal language, using lower-case symbols for variables and symbols beginning in upper-case for non-logical symbols. In defining domain relations, we will use curly angle brackets $\prec\succ$ for demarcating tuples, since we have rather overused the standard angle brackets $\langle\rangle$ already.

Definition 31 *If* π is a path with one cell—that is, $\Gamma_{U,Z5}(\pi) = \langle\langle \mathrm{U}, `CE'\rangle\rangle$ — π remains in *cell* U. If π is a path with two cells, then let $\Gamma_{U,Z5}(\pi) = \langle\langle \mathrm{U}, \mathrm{W}\rangle, \langle \mathrm{V}, \mathrm{X}\rangle\rangle$. π has an open transition *from cell* U to cell V if W *is open right and* X *is closed left.* π has a closed transition *from cell* U to cell V if W *is closed right and* X *is open left.* π has a starting transition *from cell* U to cell V if X = $`CI'$ and W *is open left.* $\pi$ has a ending transition *from cell* U to cell V if X *is open right and* W= $`CI'$.

First we define a structure over the domain of finitary paths. We begin by defining two formal languages:

The language of paths for path-connected partitions, $\mathcal{L}^p$, is the first-order language with the following predicate symbols (the arity is in parenthesis): Point(1), Cell(1), Path(1), Singleton(1), In(2), Start(2), End(2), Remains(2), ClosedTrans(3), OpenTrans(3), and Z4Splice(3), together with a collection of constant symbols U_i for $i = 1 \ldots \infty$. The language of paths for homogeneous partitions, $\mathcal{L}^c = \mathcal{L}^c \cup \{$ StartTrans(3), EndTrans(3), Z5Splice(3) $\}$.

We next define the corresponding relations over the domains of points, cells, and paths. Let **T** be a topological space and let $\mathcal{U}$ be a path-connected partition over **T**. Define the relations shown in Box 1.

(In the definitions of Z4SPLICE and Z5SPLICE above, the first vertical bar is "such that" and the second is "splice.")

For any partition $\mathcal{U}$, let $\mathcal{I}_U^p$ and $\mathcal{I}_U^c$ be the interpretations of $\mathcal{L}^p$ and $\mathcal{L}^c$ respectively mapping each predicate symbol to the corresponding relation, and mapping each symbol U_i to a cell in $\mathcal{U}$.

Second, we define a structure over ILS's. Let $\mathcal{L}^l$ be the first-order language with the following predicate symbols: Cell(1), ILS(1), Singleton(1), IRemain(2), IStartTrans(3), IEndTrans(3), ISplice(3), together with the constant symbols U_i. It is easy to show that any coherent ILS can be formed by splicing together primitive ILS's that either remain in a cell, execute a starting transition, or execute an ending transition.

We define the corresponding relations as shown in Box 2.

Let $\mathcal{I}_U^l$ be the interpretation of $\mathcal{L}^l$ mapping each predicate symbol to the corresponding relation, and mapping each symbol U_i to a cell in $\mathcal{U}$.

Third, we define a structure over gpaths in the transition graph. Let $\mathcal{L}^g$ be the first-order language with the following predicate symbols: Cell(1), GPath(1), GRemain(2), ForwardArc(3), BackwardArc(3), GSplice(3) together with the symbols U_i.

For any graph G, define the following relations:

$$\mathrm{CELLS}_G = \{\prec U \succ \mid U \text{ is a vertex in } G\}$$

$$\mathrm{GPATHS}_G = \{\prec \beta \succ \mid \beta \; is\; a\; gpath\; through\; G\}.$$

$$\mathrm{DGRAPH}_G = \mathrm{CELLS}_G \cup \mathrm{GPATHS}_G.$$

$$\mathrm{GREMAIN}_G = \{\prec \beta, \mathrm{U} \succ \mid \beta \in \mathrm{GPATHS}_G \wedge \beta = \langle \mathrm{U} \rangle\}$$

$$\mathrm{FORWARDARC}_G = \{\prec \beta, \mathrm{U}, \mathrm{V} \succ \mid \beta \in \mathrm{GPATHS}_G \wedge \beta = \langle \mathrm{U}, \langle \mathrm{U}, \mathrm{V} \rangle, \mathrm{V} \rangle\}$$

$$\mathrm{BACKWARDARC}_G = \{\prec \beta, \mathrm{U}, \mathrm{V} \succ \mid \beta \in \mathrm{GPATHS}_G \wedge \beta = \langle \mathrm{U}, \langle \mathrm{V}, \mathrm{U} \rangle, \mathrm{V} \rangle\}$$

$$\mathrm{GSPLICE}_G = \{\prec \beta_1, \beta_2, \beta_3 \succ \mid \beta_1, \beta_2, \beta_3 \in \mathrm{GPATHS}_G \wedge \beta_3 = \beta_1 \mid \beta_2\}$$

Let $\mathcal{I}^g_G$ be the interpretation of $\mathcal{L}^g$ mapping each predicate symbol to the corresponding relation, and mapping each symbol U_i to a cell in $\mathcal{U}$.

We can now state two parallel metalogical theorems. Theorem 13 states that the decision problem of a sentence in $\mathcal{L}^p$ relative to a path-connected partition can be reduced to the decision of a corresponding sentence in $\mathcal{L}^g$ relative to the transition graph. Theorem 14 states that the decision problem of a sentence in $\mathcal{L}^c$ relative to a homogeneous partition can be reduced to the decision of a corresponding sentence in $\mathcal{L}^i$ relative to the set of Z5 ILS's. In both cases the translation from the language of paths to the language of the transition graph is independent of the particular partition involved, as long as it is path-connected or homogeneous, respectively.

Theorem 13 *There exists a linear-time function* $\mathcal{A}^p$ that maps every sentence in $\mathcal{L}^p$ to a sentence in $\mathcal{L}^g$ satisfying the following. Let **T** be a Hausdorff space, let $\mathcal{U}$ be a path-connected partition over **T** with at least 2 cells, and let $\mathcal{I}^p_U$ be the interpretation of $\mathcal{L}^p$ in DPATHS_U defined above. Let G be the transition graph corresponding to $\mathcal{U}$ and let $\mathcal{I}^g_G$ be the interpretation of $\mathcal{L}^g$ in DGRAPH_G defined above, such that for each symbol U_i, $\mathcal{I}^g_G(\mathrm{U}_i) = \mathcal{I}^p_U(\mathrm{U}_i)$. Let Φ be any sentence in $\mathcal{L}^p$. Then Φ holds in the structure $\prec \mathrm{DPATHS}_U, \mathcal{L}^p, \mathcal{I}^p_U \succ$ if and only if $\mathcal{A}^p(\Phi)$ holds in the structure $\prec \mathrm{DGRAPH}_G, \mathcal{L}^g, \mathcal{I}^g_G \succ$.

Theorem 14 *There exists a linear-time function* $\mathcal{A}^c$ that maps every sentence in $\mathcal{L}^c$ to a sentence in $\mathcal{L}^l$ satisfying the following. Let **T** be a Hausdorff space, let $\mathcal{U}$ be a path-connected partition over **T** with at least 2 cells, and let $\mathcal{I}^c_U$ be the interpretation of $\mathcal{L}^c$ in DPATHS_U defined above. Let $\mathcal{I}^l_U$ be the interpretation of $\mathcal{L}^l$ in DILS_U defined above, such that for each symbol U_i, $\mathcal{I}^l_U(U_i) = \mathcal{I}^c_U(U_i)$. Let Φ be any sentence in $\mathcal{L}^c$. Then Φ holds in the structure $\prec \mathrm{DPATHS}_U, \mathcal{L}^c, \mathcal{I}^c_U \succ$ if and only if $\mathcal{A}^c(\Phi)$ holds in the structure $\prec \mathrm{DILS}_U, \mathcal{L}^l, \mathcal{I}^l_U \succ$.

The proofs are in the appendix.

These results are not actually very surprising. We have carefully crafted the languages of paths so as to exclude the expression of any information not in the transition graph, so it is no great surprise that any sentence in these languages can be translated into a sentence about the transition graph. The point of the theorems is that they give a precise characterization of what kind of information about the paths is encoded in the graph.

From the point of view of worst-case computation theory this is not actually very encouraging, as the decision problem over the graph is in fact in general undecidable.[3] However, it does give us a decision procedure for the language of paths that sometimes gives an answer, and never gives a wrong answer.

The situation as regards existential sentences (i.e. sentences with no universal quantifiers in

Box 1.

$\text{POINTS}_U = \mathbf{T}.$

$\text{CELLS}_U = \mathcal{U}$

$\text{PATHS}_U = \{\pi \mid \pi \text{ is a path through } \mathbf{T} \text{ that is finitary over } \mathcal{U}.\ \}$

DPATH_U (the domain of paths) $= \text{POINTS}_U \cup \text{CELLS}_U \cup \text{PATHS}_U$

$\text{SINGLETON}_U = \{\text{V} \mid \text{V} \in \text{CELLS}_U \wedge \exists\, \text{x} \in \text{V}\}$

$\text{IN}_U = \{\prec \text{x}, \text{V} \succ \mid \text{V} \in \text{CELLS}_U, \text{x} \in \text{V}\}$

$\text{START}_U = \{\prec \pi, \pi(0) \succ \mid \pi \in \text{PATHS}_U\}$

$\text{END}_U = \{\prec \pi, \pi(1) \succ \mid \pi \in \text{PATHS}_U\}$

$\text{REMAINS}_U = \{\prec \pi, \text{V} \succ \mid \pi \in \text{PATHS}_U \wedge \text{V} \in \text{CELLS}_U \wedge \Gamma_{U,Z5}(\pi) = \langle\langle \text{V}, \text{'CE'} \rangle\rangle\}$

$\text{CLOSEDTRANS}_U = \{\prec \pi, \text{V}, \text{W} \succ \mid \pi \in \text{PATHS}_U \wedge \text{V}, \text{W} \in \text{CELLS}_U \wedge \Gamma_{U,Z4}(\pi) = \langle\langle \text{V}, \text{'C'} \rangle, \langle \text{W}, 'R' \rangle\rangle\}$

$\text{OPENTRANS}_U = \{\prec \pi, \text{V}, \text{W} \succ \mid \pi \in \text{PATHS}_U \wedge \text{V}, \text{W} \in \text{CELLS}_U \wedge \Gamma_{U,Z4}(\pi) = \langle\langle \text{V}, 'L' \rangle, \langle \text{W}, 'C' \rangle\rangle\}$

$\text{STARTTRANS}_U = \{\prec \pi, \text{V}, \text{W} \succ \mid \pi \in \text{PATHS}_U \wedge \text{V}, \text{W} \in \text{CELLS}_U \wedge \Gamma_{U,Z5}(\pi) = \langle\langle \text{V}, 'CI' \rangle, \langle \text{W}, 'R' \rangle\rangle\}$

$\text{ENDTRANS}_U = \{\prec \pi, \text{V}, \text{W} \succ \mid \pi \in \text{PATHS}_U \wedge \text{V}, \text{W} \in \text{CELLS}_U \wedge \Gamma_{U,Z5}(\pi) = \langle\langle \text{V}, 'L' \rangle, \langle \text{W}, 'CI' \rangle\rangle\}$

$\text{Z4SPLICE}_U = \{\prec \pi, \phi, \psi \succ \mid \pi, \phi, \psi \in \text{PATHS}_U \wedge \Gamma_{U,Z4}(\psi) = \Gamma_{U,Z4}(\pi) \mid \Gamma_{U,Z4}(\phi))\}$

$\text{Z5SPLICE}_U = \{\prec \pi, \phi, \psi \succ \mid \pi, \phi, \psi \in \text{PATHS}_U \wedge \Gamma_{U,Z5}(\psi) = \Gamma_{U,Z5}(\pi) \mid \Gamma_{U,Z5}(\phi))\}$

prenex form) is more promising. Both the mappings $\mathcal{A}^p$ and $\mathcal{A}^c$ map existential sentence to existential sentences. I conjecture that the decision problem for existential sentences over the ILS structure and the transition graph structure is of the same order of computational difficulty as the word equation problem of Makanin (1977), which is known to be in PSPACE though NP-hard (Plandowski, 1999).

4. TRANSITION GRAPHS FOR SOME SAMPLE PROBLEMS

We now return from the abstract and general discussion of paths through partitioned topological spaces to the specifics of continuous spatial change. In this section we will discuss the reasoning examples enumerated in section 1 and describe the associated transition graphs, path-connected graph, and homogeneous graphs.

For our purposes the specification of a "spatial continuity problem" involves the following elements:

Box 2.

$$\begin{array}{|l|}\hline \mathrm{ILS}_U = \text{ the set of coherent Z5 ILS's over } \mathcal{U} \\ \mathrm{IREMAIN}_U = \{ \prec \alpha,\ \mathrm{V} \succ \ \mid \alpha \in \mathrm{ILS}_U \ \wedge\ \alpha = \langle\langle \mathrm{V},\ '\mathit{CE}' \rangle\rangle \} \\ \mathrm{ISTARTTRANS}_U = \{ \prec \alpha,\ \mathrm{V},\mathrm{W} \succ \ \mid \alpha \in \mathrm{ILS}_U \ \wedge\ \alpha = \langle\langle \mathrm{V},\ '\mathit{CI}' \rangle, \langle \mathrm{W},\ '\mathit{R}' \rangle\rangle \} \\ \mathrm{IENDTRANS}_U = \{ \prec \alpha,\ \mathrm{V},\mathrm{W} \succ \ \mid \alpha \in \mathrm{ILS}_U \ \wedge\ \alpha = \langle\langle \mathrm{V},\ '\mathit{L}' \rangle, \langle \mathrm{W},\ '\mathit{CI}' \rangle\rangle \} \\ \mathrm{ISPLICE}_U = \{ \prec \alpha_1,\alpha_2,\alpha_3 \succ \ \mid \alpha_1,\alpha_2,\alpha_3 \in \mathrm{ILS}_U \ \wedge\ \alpha_3 = \alpha_1 \mid \alpha_2 \} \\ \hline \end{array}$$

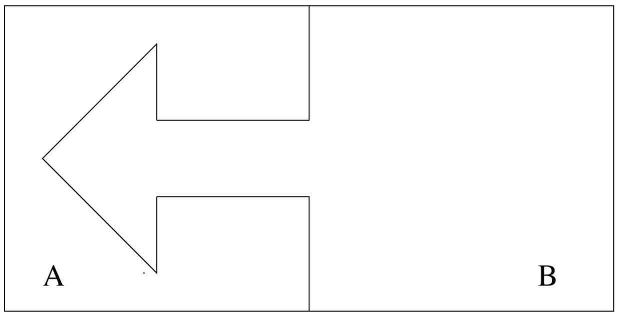
1. There are a number of *spatial fluents:* that is, entities whose value at each moment of time is a geometric entity; equivalently, fuctions from time to some space of geometric entities. In this chapter, we consider only problems with 1 or 2 non-constant spatial fluents. The state of the system at a point in time, called a *configuration*, is the tuple of the values of these fluents.
2. For each fluent, there is a specification of the general category of values it attains, such as "the class of regular open regions," "the class of points," "the class of directions," and so on. The *general configuration space* is the cross-product of these categories.
3. For each fluent, there may be additionally be a restriction on the class of values it can attain, within the general category. For instance, the region occupied by rigid object at time t is always congruent to the region it occupies at time 0. A quantity of liquid always occupies a constant volume. The *restricted configuration space* is the cross-product of these more limited classes of values.
4. There is a topology over the category of values in (2) that determines what changes are considered "continuous." Two different fluents may have values in the same category but be subject to different kinds of continuity constraints.
5. There is a JEPD (Jointly Exhaustive, Pairwise Disjoint) set of "qualitative" relations over the tuple of values of the fluents.

The space of spatial continuity problems is thus complex and diverse, and we have not found any systematic way to analyze or explore it as a whole. Lacking that, we will discuss the examples of section 1, and try to get insights into the issues involved.

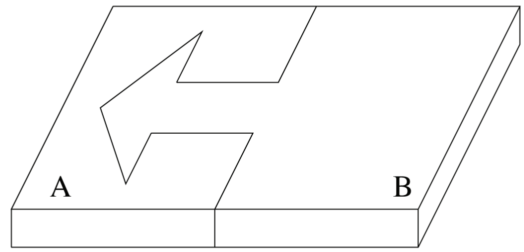
Relating this to the abstract structure developed in section 3: Let $Q_1 \ldots Q_k$ be the k spatial fluents, and let C_i be the range of values of Q_1. We assume that there is a topology defined on each C_i, as in (4) above. A point in the topological space $\mathbf{T}$ is a configuration, and the space $\mathbf{T}$ is either the general or the restricted configuration space. The topology on $\mathbf{T}$ is the cross-product of the topologies on the C_i. A path $\pi(t)$ is a tuple $\langle Q_1(t) \ldots Q_k(t) \rangle$. The partition $\mathcal{U}$ is the JEPD set of qualitative relations in (5).

A general caveat about the discussion in this section: All of the transition graphs shown below are mathematical claims and should, in principle, be proven. For many of these (e.g. the path-connected transition graph for the three-dimensional jigsaw puzzle pieces) the proof would be long, difficult, boring, and quite pointless. A complete proof requires demonstrating that each cell is homogeneous; that no coarser refinement is homogeneous; that all the arcs in the graph are possible; and that none of the arcs omitted from the graph are possible. Rather, many of these claims

Figure 6. Jigsaw puzzle pieces



Two dimensions



Three dimensions

are based on my personal geometric intuition, which is fallible. In cases where I do not feel fairly confident, I have labelled these "conjectural."

4.1. Example A: Jigsaw Puzzle Pieces

Two interlocked jigsaw puzzle pieces cannot be separated by a movement in the plane of the puzzle, but can be separated by lifting one perpendicular to the plane (see Figure 6).

Here we have an interaction of two rigid solid objects. As discussed in Section 2, this is one of the two most extensively studied problem of continuous motion in the computer science literature (the other is the problem of jointed or linked objects), almost all considering the problem of computing the configuration space from precise shape specifications.

There are a number of different ways to formalize this problem. Specifically, there are three choices with two different options each:

Two or three dimensions? As indicated in the problem statement above, the problem can be viewed in two dimensions, in which case the pieces cannot be separated, or in three dimensions, in which case they can.

Regions or placement? A state of the system can be specified, either in terms of the *regions* occupied by the two objects or in terms of their *placements*. The placement of object O at time t is a rigid mapping from some standard position of O to the position of O at t. Each of these has its advantage.

One advantage of a region-based ontology is that it generalizes easily to non-rigid entities. The flip side of this, though, is that the rigidity constraint has to be added on as an additional constraint, whereas it is built into the placement ontology.

Another advantage of a region-based ontology is that the qualitative relations—feasible, excluded—are fixed and simple relations over regions. By contrast, as functions of the placements the qualitative relations must be indexed to the underlying shape of the objects. That is, in a region-based representation, one can use a simple relation "Feasible(x,y)"; in a placement-based representation, one must use the representation "$\text{Feasible}_{\alpha,\beta}$(x,y)" where α and β are the base shapes of the objects. Viewed purely as a function of placement x and y, the latter is a strange and seemingly arbitrary region in the configuration space.

The major advantage of the placement ontology is that it gives rise to a much simpler configuration space. The configuration space of placements of a single rigid object is a three-dimensional manifold for a two-dimensional object in the plane, and a six-dimensional manifold for a three-dimensional object in three-space, and there is one standard topology. By contrast, the configuration space of regions is an strange, infinite dimensional space, with a number of different plausible topologies.

The configuration space of regions that are congruent (excluding reflection) to a reference region R_0 is isomorphic to the configuration space of placements, except when the object has some kind of symmetry, in which case multiple placements give the same region. For instance, if the three-dimensional jigsaw pieces are exactly symmetric, then the twelve homogeneous cells where the tab of B stick through the hole of A, as discussed below, reduce to 3; in the notation of Figure 10, the cells $\mathbf{A}i, \mathbf{B}i, \mathbf{C}i$ and $\mathbf{D}i$ are all identified, for each i. Though this simplifies the configuration space for symmetric objects, overall it actually tends to be an advantage of the placement ontology: the topological space of placements is constant regardless of the shape of the object, whereas the topological space of regions congruent to R_0 is isomorphic for all non-symmetric objects, but is different for symmetric objects, and each different kind of symmetry yields a different topological space. Moreover, small changes to shape that are irrelevant to the interaction—e.g. embossing the name of each jigsaw piece on its surface—will destroy the symmetry and thus alter the configuration space.

Absolute or relative position? For rigid objects, and only for rigid objects, one can view one of the objects as fixed, and characterize the position of the second relative to the first (i.e. its position in a coordinate system attached to the first).

The advantages here are closely related to those in the previous choice. A representation based on relative position does not generalize to non-rigid objects, and requires relations indexed on the shape of the fixed element or the shapes of both elements, depending on whether the ontology is region-based or placement-based. On the other hand, it reduces the dimensionality of the configuration space by a factor of 2.

In terms of our formulation of continuous spatial problems, then, this problem has the following characteristics:

Fluents: Either two fluents, for the absolute position of each object, or one fluent, for their relative position.

Category: Either the space of regular[4] bounded regions in $\mathbb{R}^k$, or the space in placements in $\mathbb{R}^k$.

Additional restriction: If the category is the space of regions, then there is the additional restriction that the region occupied by an object at any time is congruent to the region occupied at time t_0. If the category is the space of placements, then this is built in.

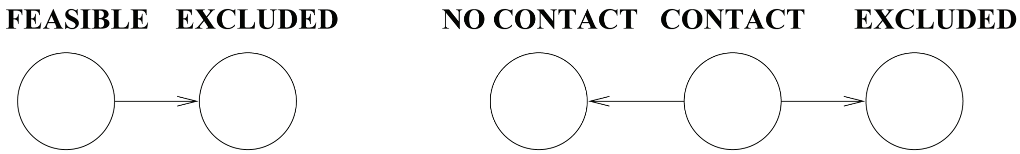
Topology. The state of placements has a single natural topology. The space of regions has several; however, they are all identical when restricted to the subspace of a rigidly moving object.

Qualitative relations: As discussed above, in a region-based representation, the starting set of qualitative relations is {**FEASIBLE, EXCLUDED** }; in a placement-based representation, it is { $\mathbf{FEASIBLE}_{\alpha,\beta}$, $\mathbf{EXCLUDED}_{\alpha,\beta}$} ,where α, β are the base shapes of the objects. Unlike Galton (1993), we will not distinguish between different categories of excluded placements, such as OV or TPP, as these make no difference in this application. If you want to categorize the relation between a solid object and an empty space which is defined by a rigid object, such as the inside of a suitcase, then OV and TPP are physically possible, and the distinction between them may be meaningful.

The simple transition graph is in Figure 7. This is the basic transition graph for the two relations {**FEASIBLE, EXCLUDED**} in almost all situations.

Whatever the shapes of the objects involved, the excluded space is always uniform in its neighbors, being open. It is generally path-connected. An exception is in the case of two-dimensional analysis of motions on a planar surface, where

Figure 7. Basic transition graphs



A: Basic transition graph

B: Basic QHD graph

the horizontal cross-sections of the obstacles that rise above the level of the floor may be disconnected (see example D below).

The feasible space for rigid objects may have multiple connected components, corresponding to different positions that cannot be attained one from another. There is always a single unbounded connected component, which includes all the configurations in which the two objects are fully separated. The remaining connected components are all bounded; there can be any number of these, or none. If object A has an interior cavity C which is large enough to hold object B, then the connected components of configuration space in which object B is in C are separate from those in which B is not in C.

In all cases, the homogeneous decomposition distinguishes between two kinds of feasible states: those that border the excluded region—i.e. those in which the two objects are in contact—and those that do not. It may also create additional distinctions, as we shall see below.

For the particular case of the jigsaw puzzle pieces: In the two dimensional case there are two connected components of feasible space: In one it is attached, and in the other it is not. The homogeneous decomposition creates the further distinction between feasible configurations where the two objects are in contact and configurations where they are separated (Figure 8)

In the three-dimensional case, the configurations where the two pieces are attached is connected in feasible space to the configurations where they are not, by sliding one of the pieces vertically. There are also configurations in which the tab of piece B goes through the hole in piece A, but I believe that, for the geometry pictured in Figure 8, these do not create additional connected components.

The homogeneous decomposition here distinguishes between configurations where the pieces are attached in the usual way but possibly with a relative vertical displacement, and those where the pieces are separable and in contact in some other way. Even though these are connected in feasible space, they differ in terms of their neighbors. The latter border the configuration region where the two pieces are separated; the former do not. The homogeneous decomposition further subdivides this into three; two limit configurations, where the top plane of piece A is aligned with the bottom plane of piece B, or vice versa; and interior configurations. The homegenous configuration also distinguishes four regions where the tab of piece B sticks through the narrow neck of piece A, meeting it on both sides (Figure 9); four, because B may face forwards or backwards and may stick downward or upward. Each of these is likewise divided into three cells—two limit regions, and the interior. The (conjectural) QHT graph is shown in Figure 10 (and Table 4).

Though this appears like a fine partition, it does not actually make all possible topological distinctions among placements. For example, state A1 includes both the configurations where the bottom face of the triangle of B is in contact with the top surface of, and those where it is not. This is another illustration of the observation made above after theorem 2, that not all points in a cell

Figure 8. Transition network for jigsaw puzzle pieces: two-dimensional motion

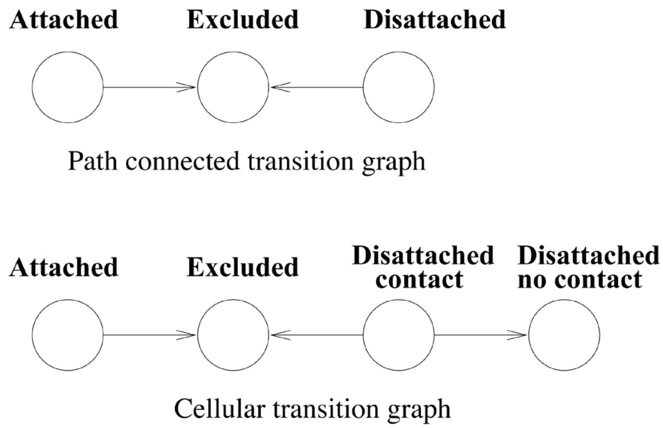
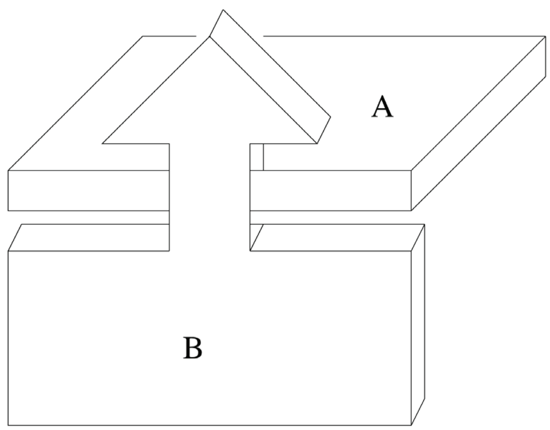


Figure 9. Jigsaw piece B sticks through A



necessarily have neighborhoods that are homeomorphic with respect to the initial partition.

4.2. Example B1: String Loop around an Hourglass

Consider a string loop of length L around the waist of an hourglass with spherical globes of circumference C. If $L > C$ then the loop can be removed from the hourglass without coming into contact with the hourglass and without ever being taut. If $L < C$ then the loop cannot be removed from the hourglass. If $L = C$, then the loop can be removed from the hourglass, but at some point it must be in contact with the hourglass, and it must be taut. It can be taken off either the upper or the lower globe.

If the globes of the hourglass are long cylinders, the circular cross section has circumference C, and $C = L$, then the string can be removed from the hourglass, but it will be taut and in contact with the hourglass over an extended interval of time.

The hourglass is a rigid object, so the issues here are the same as in example A. For simplicity, we will view the hourglass as fixed, taking its shape to be a boundary condition of the problem, and characterize the configurations of the string.

We will model the string using an idealized model in which a configuration of the string is a continuous, arc-length preserving function from the circle of circumference L into $\mathbb{R}^3$. We allow the string to cross itself, to overlay itself, or to pass through itself. A better model would additionally require that the configuration of the curve be a *simple curve*; i.e. a one-to-one function. We have not done this, because the analysis is much more complicated; the configuration space has infinitely many connected components, characterized by knot theory. If the string is taken to be of finite thickness T, then the number of connected components is exponential in L / T. Even in this simplified model, the analysis in this section should be considered conjectural.

Figure 10. QHT graph for jigsaw puzzle: three-dimensional motions

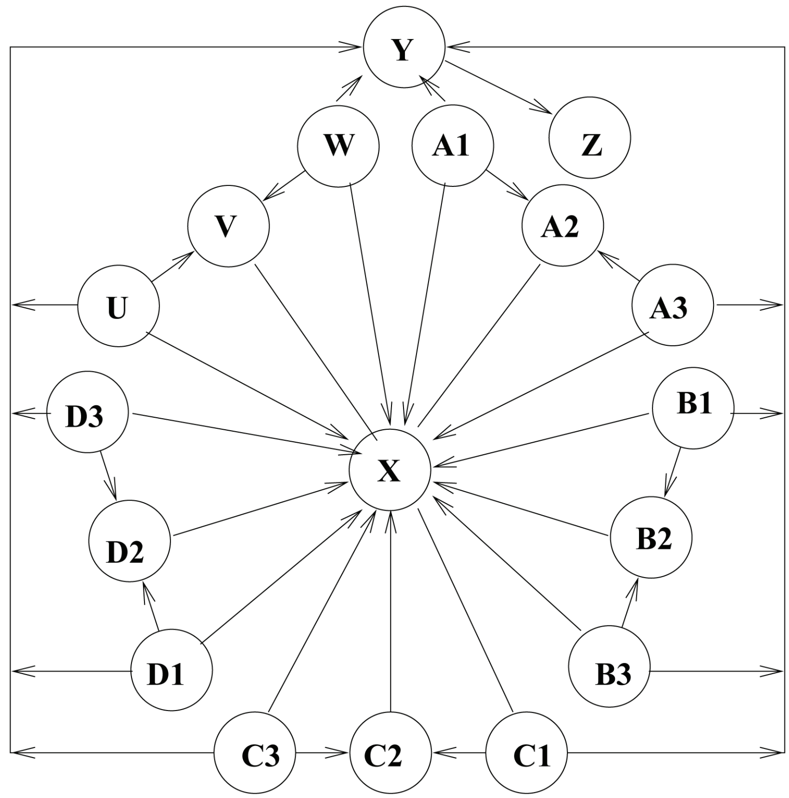


Table 4. Cell labels for Figure 10

A1: Tab of B upward through A; faces forward; borders outside.
A2: Tab of B upward through A; faces forward; interior.
A3: Tab of B upward through A; faces forward; borders inside.
B1: Tab of B upward through A; faces back; borders outside.
B2: Tab of B upward through A; faces back; interior.
B3: Tab of B upward through A; faces back; borders inside.
C1: Tab of B downward through A; faces forward; borders outside.
C2: Tab of B downward through A; faces forward; interior.
C3: Tab of B downward through A; faces forward; borders inside.
D1: Tab of B downward through A; faces back; borders outside.
D2: Tab of B downward through A; faces back; interior.
D3: Tab of B downward through A; faces back; borders inside.
U: Horizontal projections attached; top of A aligned with bottom of B.
V: Attached: no horizontal motions possible.
W: Horizontal projections attached; bottom of A aligned with top of B.
X: Excluded.
Y: Loose contact.
Z: No contact.

In terms of our format for problem specification:

Fluents: There is a single fluent, the position of the string relative to the hourglass.

Category: Let S_L be the unit circle of radius $L / 2\pi$. The category is the set of all continuous function $\phi : S_L \to \mathbb{R}^3$.

Topology: The natural metric over configurations is $d_C(\phi, \psi) = \max_{x \in S_L} d(\phi(x), \psi(x))$ where $d(a,b)$ is the usual Euclidean distance in $\mathbb{R}^3$. This is similar to the Fréchet distance between the corresponding curves; the metric here is never less and may be greater.

Additional restriction: Let A be a connected arc in S_L. Then $\phi(A)$ has a well-defined arc length, which is equal to the arc length of A.

Qualitative Relations: Let α be the region occupied by the hourglass. There are two starting qualitative regions of configuration space: $\mathbf{EXCLUDED}_\alpha$, where the string penetrates the interior of α and $\mathbf{FEASIBLE}_\alpha$ where it does not. The basic transition graph is always that shown in Figure 7a.

Let us assume that:

- The string has length L.
- The hourglass is solid; that is, we will not consider configuration in which the string lies inside the globes of the hourglass.
- The globes of the hourglass are identical spheres of radius $C / 2\pi$, with centers on the z-axis.
- The intersection of the surface of two spheres is a circle of radius $W / 2\pi$, and the hourglass is the union of the two spheres.

There are four cases to consider.

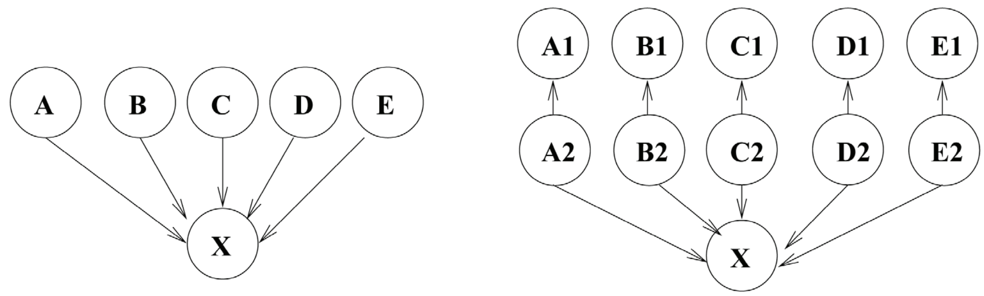
Case 1: $L < W$. The string will not go around the hourglass even at the waist. Hence, $\mathbf{FEASIBLE}_\alpha$ has a single path-connected component. The path-connected transition graph is thus still Figure 7a. The homogeneous transition graph distinguishes configurations on the boundary between $\mathbf{FEASIBLE}_\alpha$ and $\mathbf{EXCLUDED}_\alpha$; i.e. configurations where the string has contact with the surface of α (Figure 7b).

Case 2: $W \leq L < C$. Let $k = \lfloor L / W \rfloor$. Here $\mathbf{FEASIBLE}_\alpha$ has a $2k+1$ connected components, one component where it is separated from the hourglass, and $2k$ components where it is wrapped i times around the neck of the hourglass, for $i = 1 \ldots k$, clockwise or counterclockwise. The homogeneous decomposition distinguishes configurations in which the string is in contact with the hourglass from those in which it is not. If $L > kW$ then there are thus $4k+2$ cells. If $L = kW$, then there does not exist a configuration in which the string wraps k times around the hourglass and is not in contact, though there do exist two configurations in which it wraps k times and is in contact; thus there are $4k$ homogeneous cells.

Case 3: $L > C$, $(L-C) / W$ is not an integer. Let $q = \lfloor (L-C) / W \rfloor$. If the string is wrapped i times around the waist, where $q < i \leq \lfloor L / W \rfloor$, then it must remain so. As in Case 2, this gives $2(\lfloor L / W \rfloor - q)$ connected components, and either $4(\lfloor L / W \rfloor - q)$ or $4(\lfloor L / W \rfloor - q) - 2$ homogeneous cells, depending on whether L / W is an integer. If the string is wrapped q or fewer times around the waist, then it can be removed from the hourglass without coming into contact with it; all such situations form a single connected component with the disattached configurations, and form two homogeneous cells (contact and non-contact).

Case 4: $C + qW = L$ for integer $q \geq 0$. This is the same as Case 3, with the following change: If the string is wrapped $q+1$ times around the waist, then it is possible to separate the string from the hourglass, taking off one loop at a time; however, there is necessarily an instant at which the string is taut around both the globe and the waist. These configurations are thus also part of the connected component with the disattached

Figure 11. Transition graphs for string and hourglass: cases 2 and 3



Path connected graph: Case 2–4   Homogeneous graph: Case 2,3

configurations, but give rise to ten homogeneous cells: wrapped around the waist, contact or no contact, clockwise or counter-clockwise (4 cells); taut around the upper/lower globe, clockwise and counter-clockwise (4 cells); and disattached, contact or no contact (2 cells).

If the globes of the hourglass are cylinders whose central axis is longer than $L / 2$ (the geometry for short cylinders is hairy), then the string can be removed at all from a position wrapped i times around the waist if and only if $L \geq iC$. So, if $i > L / C$ it cannot be removed; this gives $2(\lfloor L / W \rfloor - \lfloor L / C \rfloor)$ connected components (each value of i, clockwise and counterclockwise) and $4(\lfloor L / W \rfloor - \lfloor L / C \rfloor)$ homogeneous cells (contact and no contact, unless $i = L / W$ is an integer, as in case 2). If $i < L / C$ then it can be removed without contact; these configurations are therefore all part of the same single component that includes disattached configurations. This component has two homogeneous cells. If $i = L / C$ is an integer, then this is also part of the disattached connected components, but gives rise to eighteen homogeneous cells: each of the taut cells of Case 4 above is split into three (at inner rim, at outer rim, in middle).

Figures 11 and 12 show the transition graphs for these four cases as explained in Table 5

4.3. Example B2: Rubber Band around an Hourglass

If, instead of a string loop, we have a rubber band whose length is less than C at rest but can be stretched to a length greater than C, then it can be removed from the hourglass without being in contact with the hourglass, but it must be stretched in order to do so.

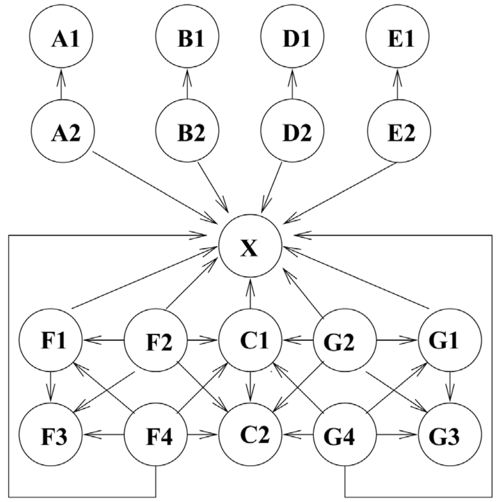
The problem specification for this example differs from Example B1 in two respects: First, the **Additional restrictions** becomes the following: Let A be a connected arc in S_L. Then $\phi(A)$ has a well-defined arc length $|\phi(A)|$ and $|A| \leq \phi(A) \leq \gamma |A|$ where γ is the ratio between the length of the band when maximally stretched and the length of the band when relaxed.

Second, the *Qualitative relations* involve a distinction between whether the band is relaxed or stretched . Thus, the basic set of qualitative relations has four elements:

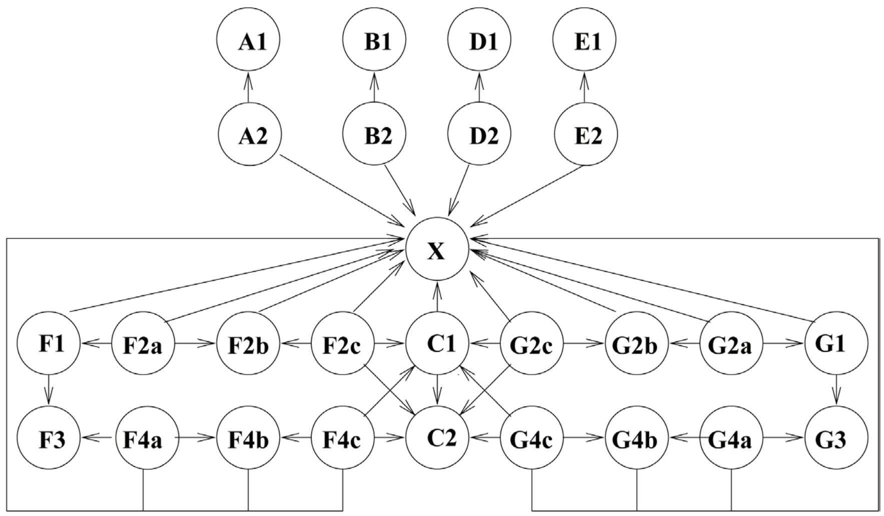
$\{$**FEASIBLE**$_\alpha \wedge$ **RELAXED**,
FEASIBLE$_\alpha \wedge$ **STRETCHED**,
EXCLUDED$_\alpha \wedge$ **RELAXED**,
EXCLUDED$_\alpha \wedge$ **STRETCHED**$\}$

The transition graphs are formed by combining the transition graph from Example B1 for a string whose length is the relaxed length of the rubber

Figure 12. Transition graphs for string and hourglass: case 4 and cylinder



Homogeneous graph: Case 4: spheres



Homogeneous graph: cylinders

band, and the transition graph for a string whose length is the maximal length of the rubber band.

4.4. Example C: Milk in a Bottle

A quantity of milk in a closed bottle remains in the bottle.

There are a number of different says to characterize the continuous motion of a liquid (Davis, 2008, 2009). For our purposes here, the simplest and most relevant is as follows: The fluent is the region occupied by the milk, which we take to be a regular region. Since liquid is incompressible—that is, a body of liquid occupies a constant

Table 5. Transition graphs for string and hourglass: key

	Case 2: $L = 2.5W < C$.	Case 3: $C = 2.5W, L = 4.2W$.
Path-connected graph		
A:	Wrapped twice CW	Wrapped four times CW
B:	Wrapped once CW	Wrapped three times CW
C:	Detachable	Detachable
D:	Wrapped twice CCW	Wrapped four times CCW
E:	Wrapped once CCW	Wrapped three times CCW
X:	Excluded	Excluded
homogeneous graph		
A1:	Wrapped twice CW, contact	Wrapped four times CW, contact
A2:	Wrapped twice CW, no contact	Wrapped four times CW, no contact
B1:	Wrapped once CW, contact	Wrapped three times CW, contact
B2:	Wrapped once CW, no contact	Wrapped three times CW, no contact
C1:	Detachable, contact	Detachable, contact
C2:	Detachable, no contact	Detachable, no contact
D1:	Wrapped twice CCW, contact	Wrapped four times CCW, contact
D2:	Wrapped twice CCW, no contact	Wrapped four times CCW, no contact
E1:	Wrapped once CCW, contact	Wrapped three times CCW, contact
E2:	Wrapped once CCW, no contact	Wrapped three times CCW, no contact
Case 4: $C = 2.5W, L = 4.5W$.		
A,B,C,D,E,X,A1,A2,B1,B2,D1,D2,E1,E2: as in case 3.		
C1: Detached, contact.		
C2: Detached, no contact.		
F1: Wrapped twice around waist CW, contact.		
F2: Taut around upper globe and waist, CW.		
F3: Wrapped twice around waist CW, no contact.		
F4: Taut around lower globe and waist, CW.		
G1: Wrapped twice around waist CCW, contact.		
G2: Taut around upper globe and waist, CCW.		
G3: Wrapped twice around waist CCW, no contact.		

continued on following page

Table 5. Continued

G4: Taut around lower globe and waist, CCW.
Cylindrical globes, $L = 4.5W = 2C$.
A,B,C,D,E,X,A1,A2,B1,B2,C1,C2,D1,D2,E1,E2,F1,F3,G1,G3: as in case 4.
F2a: Taut around cylinder CW, at outer rim, upper cylinder.
F2b: Taut around cylinder CW, not at rim, upper cylinder.
F2c: Taut around cylinder CW, at inner rim, upper cylinder.
F4a: Taut around cylinder CW, at outer rim, lower cylinder.
F4b: Taut around cylinder CW, not at rim, lower cylinder.
F4c: Taut around cylinder CW, at inner rim, lower cylinder.
G2a: Taut around cylinder CCW, at outer rim, upper cylinder.
G2b: Taut around cylinder CCW, not at rim, upper cylinder.
G2c: Taut around cylinder CCW, at inner rim, upper cylinder.
G4a: Taut around cylinder CCW, at outer rim, lower cylinder.
G4b: Taut around cylinder CCW, not at rim, lower cylinder.
G4c: Taut around cylinder CCW, at inner rim, lower cylinder.

volume—a natural metric to use is the following (original here, to the best of my knowledge):

Definition 32 *Let V(*R*)* be the volume of region R . A function Φ over $R1$ is volume-preserving if, for every subset $R \subset R1$, V($\Phi(R)$) = V(R). The cost of Φ on $R1$, $C(\Phi, R1) = \sup_{p \in R} d(p, \Phi(p))$.

Let $R1$ and $R2$ be closed, bounded, regular regions such that V($R1$)=V($R2$). The volume-preserving distance from $R1$ to $R2$, $d_{vp}(R1, R2)$ is the infimum of $C(\Phi, R1)$ for all volume-preserving Φ such that Cl($\Phi(R1)$)= $R2$ (Note that Φ need not be continuous, and therefore $\Phi(R1)$ may not be closed).

Example (Figure 13a): Let $R1$ be the unit square [0,1]$\times$[0,1], and let $R2$ be the pair of rectangles [0,1/2]$\times$[0,1] $\cup$ [3/2,2]$\times$[0,1]. Let Φ be the function:

$$\Phi(p) = \begin{cases} p & if \ p_x \leq 1/2 \\ p + \hat{x} & if \ p_x > 1/2 \end{cases}$$

Then C$(\Phi, R1) = 1$. It is easily shown that no volume-preserving function has lower cost, so $d_{vp}(R1, R2) = 1$.

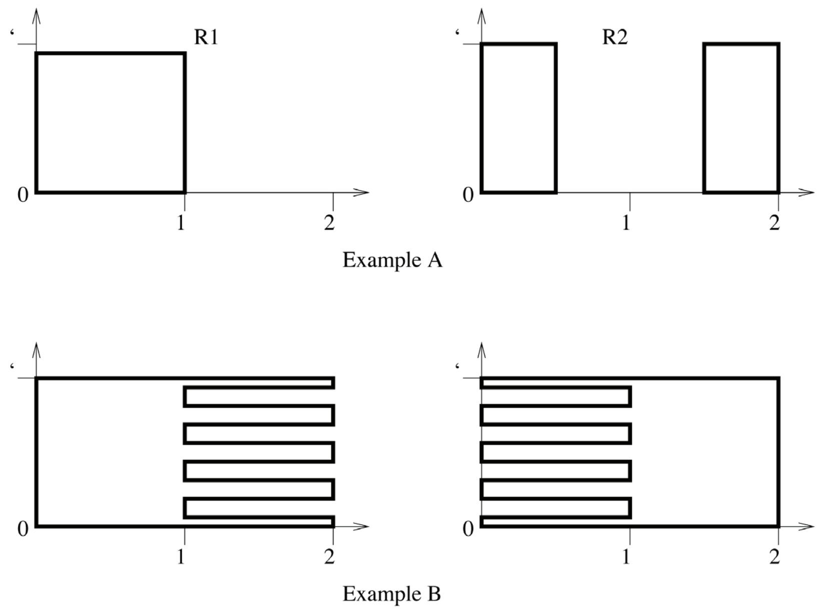
It is easily shown that d_{vp} is a metric over closed regular regions, and that it is always at least as large as the Hausdorff distance, and often greater. For instance in Figure13b, the Hausdorff distance from $R1$ to $R2$ is 1/16, since every point in $R1$ is within distance 1/16 of a point in $R2$ and vice versa. A volume-preserving function, however, must move one square unit of liquid from the left-hand side to the right hand side; it can be shown that $d_{vp}(R1, R2) = 1/2$. Taking this kind of example to the limit, it follows that the topology induced by d_{vp} is stricly finer than that induced by the Hausdorff distance.

In terms of our format for problem specification:

Fluent: There is a single fluent, which is the region occupied by the milk. We take the region α occupied by the bottle to be a constant boundary condition.

Category: The value of the fluent is a regular region.

Figure 13. Volume preserving distance



Additional constraint: The volume of the fluent is a constant value V .

Topology: The topology defined by the metric d_{vp} .

Qualitative relation: EXCLUDED$_\alpha$, in configurations where the milk penetrates the interior of the bottle, and **FEASIBLE**$_\alpha$ where it does not.

The basic transition graph is the standard one (Figure 7).

The feasible region actually has uncountably many connected components: for each $V1$ between 0 and V inclusive, the set of configurations with volume $V1$ of milk inside the bottle and $V - V1$ outside is a separate connected component. The effect of a homogeneous decomposition is just to divide each of these connected components into two parts: configurations where the milk is in contact with the bottle (either on the inside or outside), and those where it is not (Figure 14).

Nonetheless, the partition is locally finite, so the theory developed in Section 3 still applies.

4.5. Example C2: Milk in Cups

If at time $T1$ there is milk sitting in open cup A, and at a later time $T2$ the milk has moved to a cup B, and both cups are stationary, then the milk came out of the top of cup A and went in the top of cup B.

The problem formulation is the same as in C1, except that the shape of the cups is different from the shape of the bottle, and that the basic qualitative relations are different. Specifically, we defined a region RC as *cupped* by region RO if the boundary of RC is the union of two parts, $BO \cup BT$ where BO is a subset of the boundary of RO , and BT is a surface lying in a horizontal plane above RC (Figure 15). We then consider a basic set of four possible qualitative relations:

- **EXCLUDED**: configurations where the milk penetrates the interior of the material of cup.
- **A** if all of the milk is in a region cupped by cup A.

Figure 14. Transition network for milk and bottle

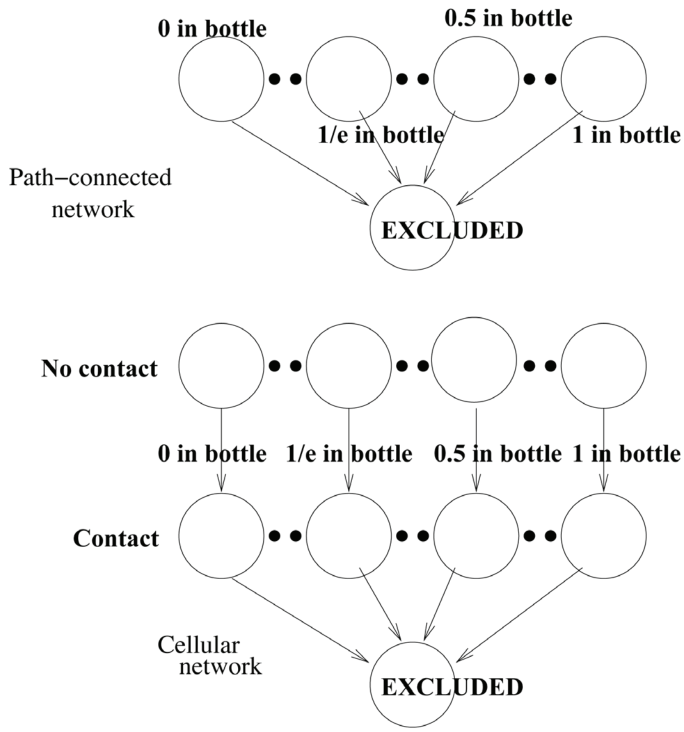
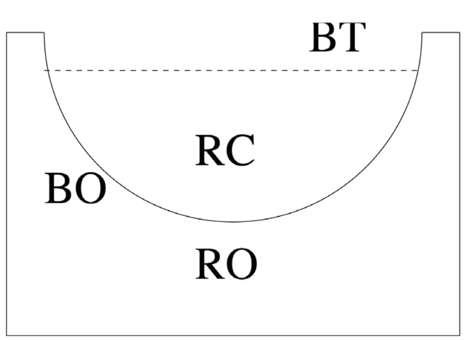


Figure 15. Cupped region



- **B** if all of the milk is in a region cupped by cup B.
- **OTHER** if the configuration is feasible, but neither **A** nor **B** holds.

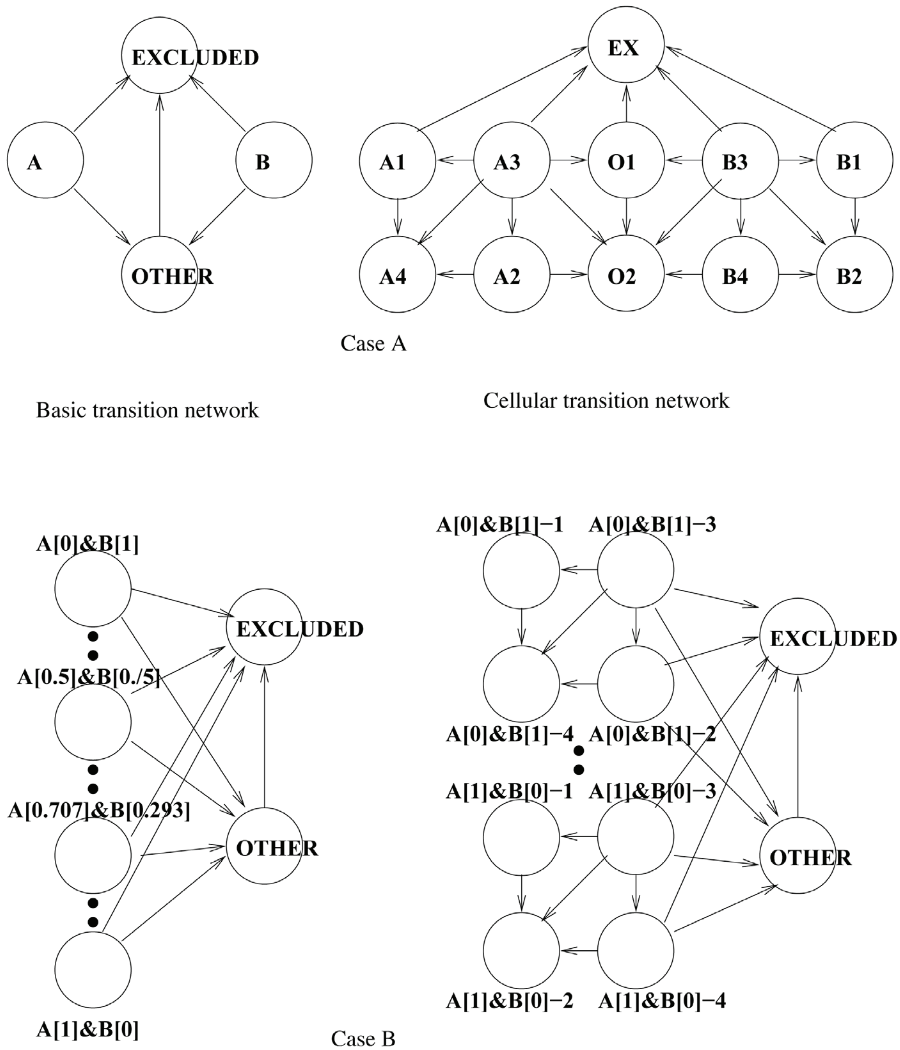
Each of these cells is path-connected. A homogeneous decomposition divides cells **A** and **B** each into four subcells, according to whether the milk is (**A1**) in contact with the cup, (**A2**) with the opening at the top, (**A3**) both, or (**A4**) neither. It divides **OTHER** into two subcells, according to whether the milk is (**O1**) in contact with either cup or (**O2**) not (Figure 16a).

If we modify the problem to read that some but not all of the milk moves from cup A to cup B, then that can be formalized using a starting partition that has uncountably many cells, though it is locally finite:

- **EXCLUDED:** The milk penetrates the material of the cup.
- $\mathbf{A}[\mathrm{V1}]$ & $\mathbf{B}[\mathrm{V}-\mathrm{V1}]$: For any $V1$ between 0 and V, there is volume $V1$ in cup A and $V-V1$ in cup B.
- **OTHER**: The configuration is feasible and some of the milk is outside both cups.

Each of these is path-connected. The homogeneous decomposition divides the cupped com-

Figure 16. Transition networks for milk in cups



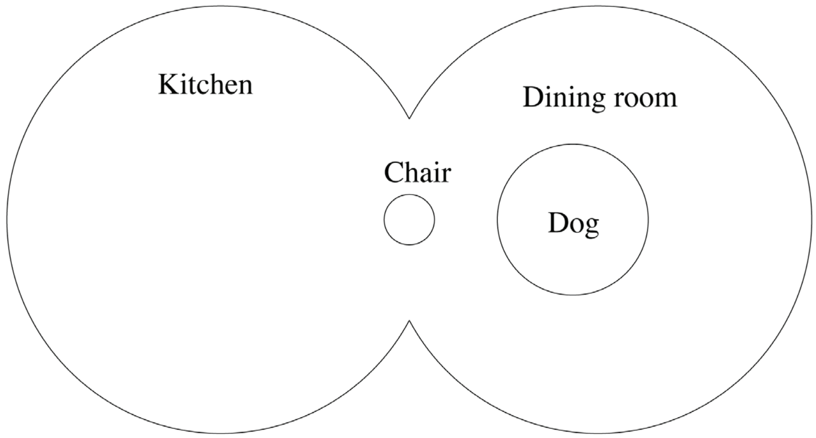
ponents into four, and the **OTHER** component into two, as in the earlier analysis (Figure 16b).

4.6. Example D: Blocking the Dog with a Chair

The dog can go from the dining room into the kitchen. However, if a chair is placed in the middle of the kitchen doorway, then the dog cannot go from the dining room to the kitchen. If the chair is placed at the edge of the doorway, then the dog can squeeze past and get into the kitchen.

The transition graph analysis of this problem is problematic. The problem is that if a configuration is taken to be the pair of the position of the chair and the position of the dog, then the class of paths includes any simultaneous motion of the dog and the chair. The transition graph over this set of paths cannot represent the above conclusions, which describe the possible motions of the dog in which the chair is static. In fact, the homogeneous transition graph just has the three states **NO CONTACT, CONTACT,** and **EXCLUDED** (Figure 7b).

Figure 17. Dog and chair scenario



We can solve this problem by using a hierarchy of transition graphs, with two layers. The bottom layer of transition graphs describes the motions of the dog for possible fixed positions of the chair. The upper layer describes the motions of the chair, characterized by the the transition graphs of the lower level. That is, we take the starting qualitative relations over the position of the chair to be the transition graphs from the lower level (More precisely, equivalence classes of transition graphs under isomorphism; the graphs for any two positions of the chair are, strictly speaking, not equal because the corresponding cells created for the dog are not actually equal, just close.).

To simplify the analysis, I have taken the dog, the chair, and the walls of the two rooms to be circular, as shown in Figure 17; more realistic shapes will lead to both more complex transition graphs and more difficult analysis. Even with this simplification, the analysis here of the homogeneous case is somewhat conjectural. Let Dist(C,W) be the distance from the chair to the wall, and let Diam(D) be the diameter of the dog. The width of the doorway is equal to Dist(C,W)+Diam(D).

In a layered representation using path connected graphs, there are three lower-level transition graphs:

A. The chair does not block the dog, and Dist(C,W) < Diam(D). In this case there are two cells: (**A1**) feasible and (**A2**) excluded.
B. The chair does not block the dog, and Dist(C,W) ≥ Diam(D). In this case, the excluded region has two connected components, so there are three cells: (**B1**) feasible; (**B2**) dog overlaps walls; (**B3**) dog overlaps chair.
C. The chair blocks the dog. In this case the feasible region has two connected components so there are three cells: (**C1**) dog in dining room; (**C2**) dog in kitchen; (**C3**) excluded.

The upper-level transition graph has four cells (Figure 18):

U1: The chair does not block the dog, and Dist(C,W) < Diam(D). Lower-level graph A.
U2: The chair does not block the dog, Dist(C,W) ≥ Diam(D), and the chair is in the dining room. Lower-level graph B.
U3: The chair does not block the dog, Dist(C,W) ≥ Diam(D), and the chair is in the kitchen. Lower-level graph B.
U4: The chair blocks the dog. Lower-level graph C.

Figure 18. Layered path-connected transition network for dog and chair

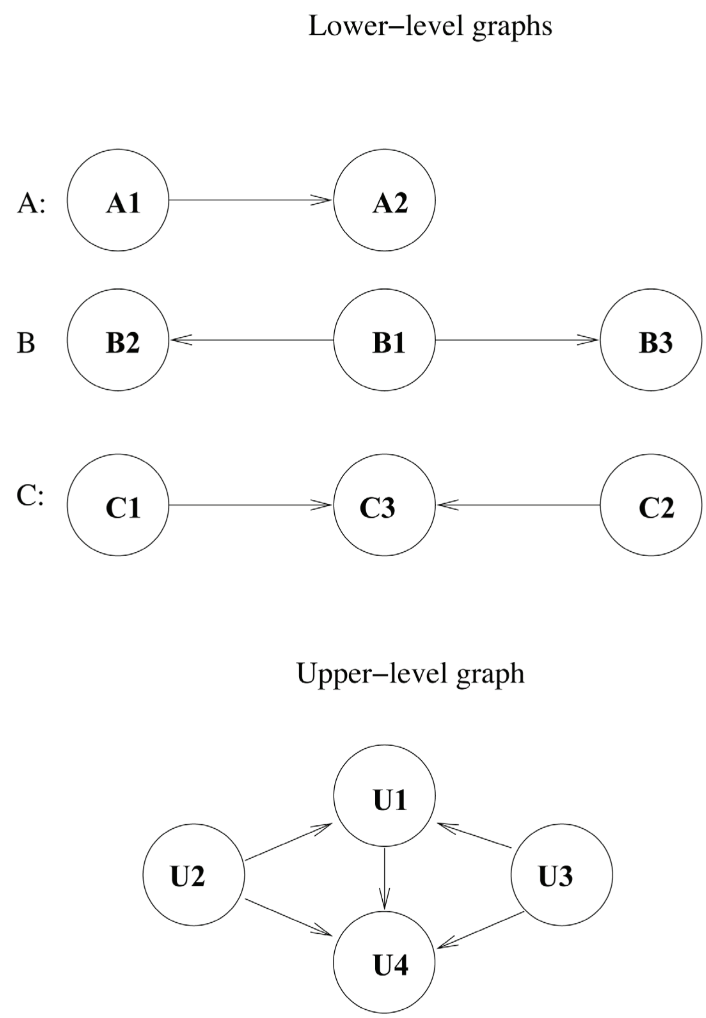
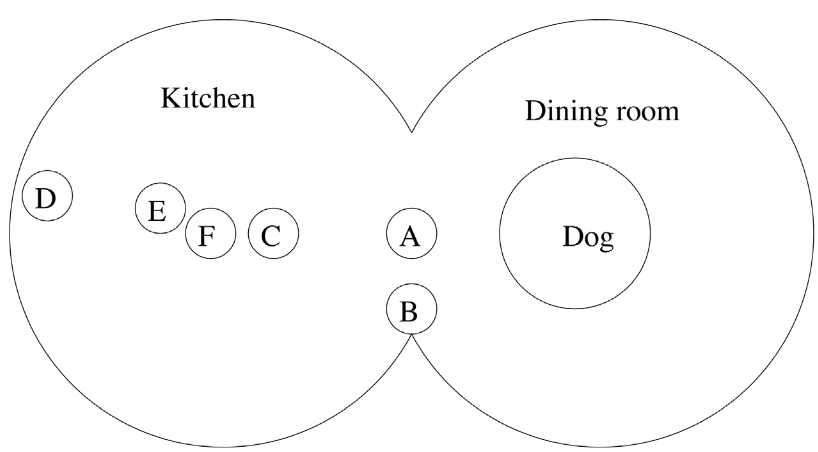


Figure 19. Positions of chairs creating different low-level homogeneous graphs



In a layered representation using homogeneous graphs, there are six transition lower-level transition graphs (Figures 19 and 20).

A. The chair blocks the dog. The homogeneous graph here has five cells: Dog in kitchen, (**A1**) with or (**A2**) without contact; dog in dining room,(**A3**) with or (**A4**) without contact; (**A5**) excluded.
B. The chair in the doorway allows the dog to squeeze by on one side. The homogeneous graph has six cells: Dog in kitchen, (**B1**) with or (**B2**) without contact; dog in dining room, (**B3**) with or (**B4**) without contact; (**B5**) dog squeezing by; (**B6**) excluded.
C. The chair is placed so as to allow the dog to squeeze by on either side. The homogeneous graph has ten cells: Dog in kitchen, (**C1**) with or (**C2**) without contact; dog in dining room, (**C3**) with or (**C4**) without contact; dog squeezing by (**C5**) on left or (**C6**) on right; dog in contact with chair (**C7**) on the kitchen side or (**C8**) on the dining room side; excluded because (**C9**) dog overlaps walls; (**C10**) dog overlaps chair.
D. The chair does not block the dog, and Dist(C,W) $<$ Diam(D). The homogeneous graph has three states: (**D1**) dog not in contact, (**D2**) dog in contact, (**D3**) excluded.
E. The chair does not block the dog, and Dist(C,W)=Diam(D). The homogeneous graph has six states: (**E1**) dog not in contact, (**E2**) dog in contact with walls, (**E3**) dog in contact with chair, (**E4**) dog in contact with both walls and chair, (**E5**) dog overlaps walls, (**E6**) dog overlaps chair.
F. The chair does not block the dog, and Dist(C,W) $>$ Diam(D). The homogeneous graph has five states: (**F1**) dog not in contact, (**F2**) dog in contact with walls, (**F3**) dog in contact with chair; (**F4**) dog overlaps walls, (**F5**) dogs overlaps chair.

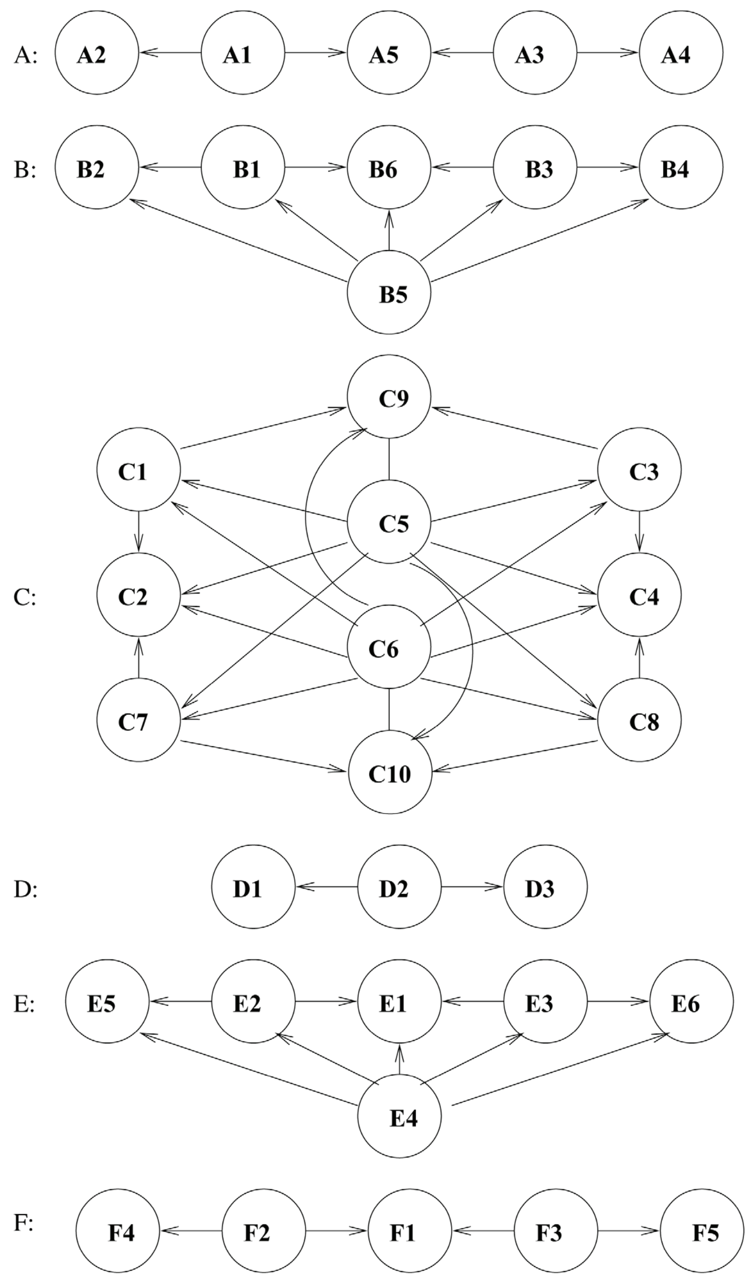
The upper-level transition graph has eighteen cells (Figure 21):

U1a – U4b: Chair does not interfere with dog. Cross product of { Dist(C,W)=0 / Dist(C,W) $<$ Diam(D) / Dist(C,W) $=$ Diam(D) / Dist(C,W) $>$ Diam(D) } times { Chair in kitchen, chair in dining room }. Lower-level graphs: D, D, D, D, E, E, F, F respectively.
U5: Chair blocks dog. Lower-level graph A.
U6: Dog can squeeze through on right. Chair is (**U6a**) in dining room, (**U6b**) in kitchen, (**U6c**) against the left wall. Lower-level graph B.
U7: Dog can squeeze through on left. Chair is (**U7a**) in dining room, (**U7b**) in kitchen, (**U7c**) against the right wall. Lower-level graph B.
U8: Dog can squeeze through on both sides. Chair is (**U8a**) in dining room, (**U8b**) in kitchen. Lower-level graph C.
U9: Excluded: Chair overlaps walls. Lower-level graph D.

In terms of our structure for problem specifications: At the lower level, the fluent is the position of the dog. The set of qualitative relations is $\{$ **EXCLUDED**$_\alpha$, **FEASIBLE**$_\alpha\}$ where α is the region occupied by the walls and by the chair. At the upper level, the fluent is the position of the chair. The starting set of qualitative relations is the set of transition networks from the lower level, plus the excluded state. The remainder of both problem is the same as in example A.

However, this analysis is not satisfactory, because it introduces a false assymetry between the chair and the dog. It is equally true, after all, that if the dog has settled down in the center of the doorway, then one cannot move the chair from one room to another, but this transition graph structure expresses that only very indirectly.

Figure 20. Lower-level homogeneous graphs for dog and chair



4.7. Example E: Travel from Alaska to Idaho

A person who is in Canada at one time and in the United States at a later time must cross the U.S. border at some time in between. A person who is in Alaska at one time and in Idaho at a later time must cross the U.S. border at least twice in between. It is possible to travel from any point in Idaho to any point in Ohio without crossing the border of the United States.

This is actually the simplest of our examples in terms of the configuration space. However, it is the most complex in terms of the qualitative rela-

Figure 21. Upper-level transition graph for dog and chair

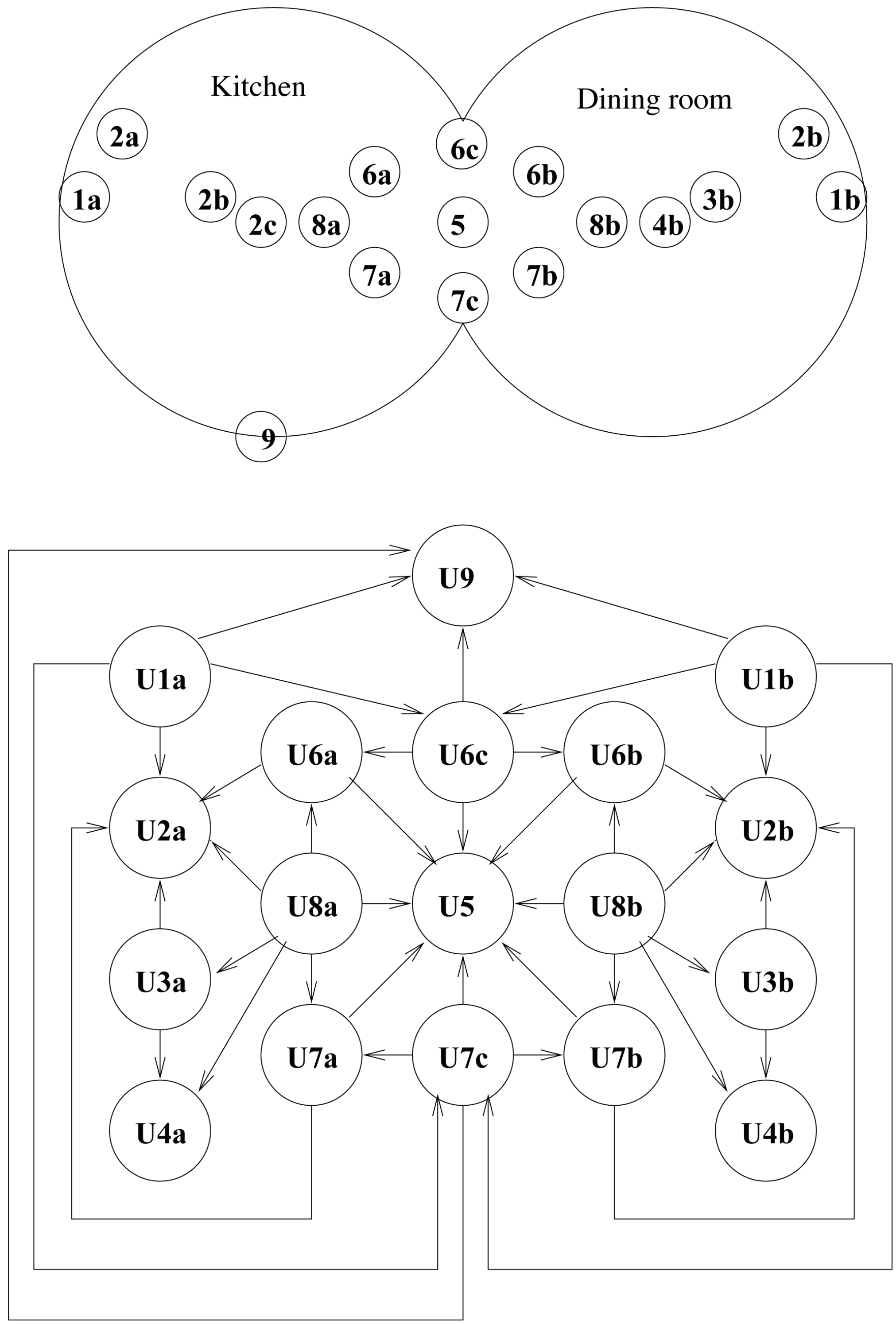
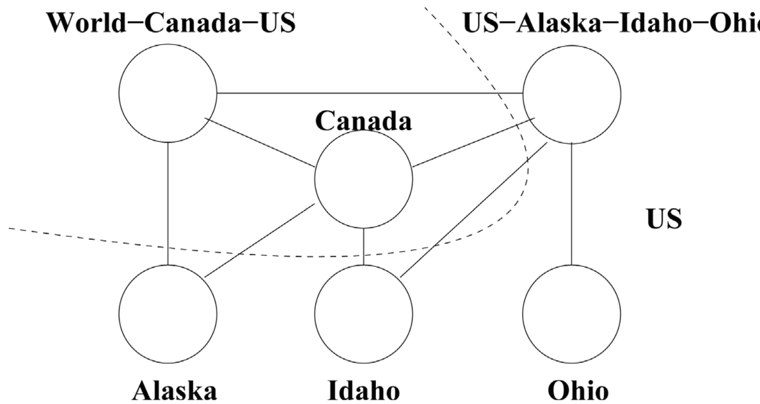


Figure 22. Transition network for traveller



tions and consequently the least suited to analysis in terms of our theory of transition graphs.

We idealize a person as a point object, which is reasonable in comparison to Alaska. There is a single fluent, which is the person's position. The configuration space is the surface of the earth's sphere, under the usual topology. There are no additional restrictions.

However, qualitative relations such as "in Alaska," "in the United States," and so on, do not form a JEPD set; they are neither exhaustive nor disjoint. Rather, these are a set of regions in configuration space that are related by RCC relations. Therefore, this problem cannot be characterized in terms of transition graphs in anything like the form developed in this chapter.

One could turn these into a JEPD set by constructing all the non-empty complete Boolean combinations of the base relations, analogous to a frame of discrimination in probabilistic reasoning. The base relations are then characterized as a set of atomic relations. Figure 22 shows the transition graph for the particular regions in the above problem. However, this is not a very satisfactory solution. First, even if a complete characterization of the RCC relations between the given regions is given, that does not at all determine the relations between the Boolean combinations, still less the structure of the path-connected and homogeneous refinements. For instance, the first step in finding the path-connected refinement of Figure 22 is the very difficult one of determining how many path-connected components the US and Canada have. Second, the set of relations can become exponentially large, and reasoning about the base relations involves reasoning about these large sets. Third, the set of relations is likely to be overly fine for the given application.

5. CONCLUSION AND FUTURE WORK

In this chapter, we have defined the transition graph for a JEPD set of relations over a configuration space, and we have defined two refinements of the basic graph: the graph of path-connected components and the homogeneous graph. We have proven metalogical theorems stating that the decision problem for specified first-order languages over continuous paths through configuration space is reducible to the decision problem for paths in the path-connected and homogeneous transition graphs; and that the same holds for the existential subset of those languages. We have shown how these techniques can be applied in a range of physical reasoning problems.

Many problems remain to be solved, however. The most important problem is the problem of deriving these transition networks from the problem specification. Currently, the only case where this is a well-developed theory for this is for the case of rigid objects with exact shape specification; this is the piano-movers problem. In a broader setting, such as the problems discussed in this chapter, we do not even have a reasonable representation language for problem specification, let alone an algorithm for deriving a transition network.

Other important problems include finding better techniques for solving problems like examples D and E of this chapter. In example D, there are several moving objects, and one wants to reason about moving one of these at a time. In example E, the qualitative relations involved are not JEPD; their topological relations are known, possibly incompletely.

In general, the examples in this paper suggest that qualitative homogeneous decompositions are not qualitative enough; they tend to introduce large number of distinctions that are of no actual value for the applications involved.

REFERENCES

Collins, G. E. (1975). Lecture Notes in Computer Science: *Vol. 33*. *Quantifier elimination for real closed fields by cylindrical algebraic decomposition* (pp. 134–183). Berlin, Germany: Springer-Verlag.

Davis, E. (1990). *Representations of commonsense knowledge*. San Mateo, CA: Morgan Kaufmann.

Davis, E. (2000). *Describing spatial transitions using mereotopological relations over histories. Technical Report #2000-809*. New York, NY: New York University.

Davis, E. (2001). Continuous shape transformations and metrics on regions. *Fundamenta Informaticae*, *46*(1-2), 31–54.

Davis, E. (2008). Pouring liquids: A study in commonsense physical reasoning. *Artificial Intelligence*, *172*, 1540–1578. doi:10.1016/j.artint.2008.04.003

Davis, E. (2009). Ontologies and representations of matter. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*. AAAI.

de Kleer, J. (1977). Multiple representations of knowledge in a mechanics problem solver. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77)*, (pp. 299-304). IJCAI.

de Kleer, J., & Brown, J. S. (1985). A qualitative physics based on confluences . In Bobrow, D. (Ed.), *Qualitative Reasoning about Physical Systems* (pp. 7–83). Cambridge, MA: MIT Press.

Forbus, K. (1980). Spatial and qualitative aspects of reasoning about motion. In *Proceedings of the First National Conference on Artificial Intelligence, (AAAI-80)*, (pp. 170-173). AAAI.

Forbus, K. (1985). Qualitative process theory . In Bobrow, D. (Ed.), *Qualitative Reasoning about Physical Systems* (pp. 85–186). Cambridge, MA: MIT Press.

Galton, A. P. (1993). Towards an integrated logic of space, time, and motion. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, (IJCAI-93)*, (pp. 1550-1555). San Mateo, CA: Morgan Kaufmann.

Galton, A. P. (1995). Towards a qualitative theory of motion. In [Berlin, Germany: Springer-Verlag.]. *Proceedings of*, *COSIT-95*, 377–396.

Galton, A. P. (2000a). *Qualitative spatial change*. Oxford, UK: Oxford University Press.

Galton, A. P. (2000b). Continuous motion in discrete space. In *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning,* (pp. 26-37). Knowledge Representation and Reasoning.

Galton, A. P. (2003). A generalized topological view of motion in discrete space. *Theoretical Computer Science*, *305*, 111–134. doi:10.1016/S0304-3975(02)00701-6

Hayes, P. (1979). The naive physics manifesto . In Michie, D. (Ed.), *Expert Systems in the Microelectronic Age* (pp. 242–270). Edinburgh, UK: Edinburgh University Press.

Hazarika, S., & Cohn, A. G. (2011). *Qualitative spatial change in a mereotopology*. Unpublished.

Kuipers, B. (1986). Qualitative simulation. *Artificial Intelligence*, *29*, 289–338. doi:10.1016/0004-3702(86)90073-1

LaValle, S. (2006). *Planning algorithms*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511546877

Lozano-Perez, T. (1983). Spatial planning: A configuration space approach. *IEEE Transactions on Computers*, *32*(2), 108–120. doi:10.1109/TC.1983.1676196

Makanin, G. (1977). The problem of solvability of equations in a free semigroup. *Matematicheskii Sbornik*, *32*, 129–198. doi:10.1070/SM1977v-032n02ABEH002376

Muller, P. (1998). A qualitative theory of motion based on spatio-temporal primitives. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR-98),* (pp. 131-141). San Mateo, CA: Morgan Kaufmann.

Muller, P. (1998b). *Éléments d'une théorie du mouvement pour la modélisation du raisonnement spatio-temporel de sens commun.* Doctoral Dissertation. Toulouse, France: Université Paul Sabatier.

Plandowski, W. (1999). Satisfiability of word equations with constants is in PSPACE. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS '99),* (pp. 495-500). Los Alamitos, CA: IEEE Computer Society Press.

Randell, D. A., & Cohn, A. G. (1989). Modelling topological and metrical properites of physical processes. In *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning, (KR-89),* (pp. 55-66). KR.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning, (KR-92),* (pp. 165-176). KR.

Weidenbach, C. (2001). SPASS: Combining superposition, sorts, and splitting . In Robinson, A., & Voronkov, A. (Eds.), *Handbook of Automated Reasoning* (*Vol. 2*, pp. 1965–2013). Amsterdam, The Netherlands: Elsevier.

Williams, B. (1985). Qualitative analysis of MOS circuits . In Bobrow, D. (Ed.), *Qualitative Reasoning about Physical Systems* (pp. 281–346). Cambridge, MA: MIT Press.

Key Terms and Definitions

Configuration Space: A point in a configuration space corresponds to one spatial arrangement of the objects under consideration.

First-Order Definable: Structure S is first-order definable in structure T if a collection of relations C isomorphic to the relations in S can be defined in the first-order language of T.

First-Order Equivalent: Two structures with the same language are first-order equivalent if the same first-order sentences are true in both.

Partition: P is a partition of T if P is a collection of subsets of T and every element of T is in exactly one set in P.

Path: A path in a space is a continuous function from a closed time interval into the space.

Path-connected: A region R is path-connected if every pair of points in R is connected by a path that remains in R.

RCC-8: A exhaustive: disjoint, collection of topological relations between pairs of regions in Euclidean space. Two regions may be equal (EQ), overlap (OV), externally connected (EC), disconnected (DC), one may be a tangential partial part of the other (TPP and TPP^{-1}), or one may be a non-tangential partial part of the other (NTPP and $NTPP^{-1}$) [Randell, Cui, and Cohn, 1992].

Transition Graph: A directed graphs whose vertices are regions of configuration space and whose edges represent a possible transition.

ENDNOTES

1. It is not known whether this stronger definition corresponds to continuity relative to any topology over the space of regular regions. I would conjecture that it does not. It is certainly strictly stronger than continuity with respect to the Hausdorff distance and strictly weaker than continuity with respect to the dual-Hausdorff distance (Davis, 2001). Even if the conjecture is true, that does not indicate that the definition is flawed, just that standard results about continuity do not necessarily apply.
2. A Hausdorff topology satisfies the constraint that for any points **x,y** there exists disjoint open sets U, V such that $\mathbf{x} \in U$ and $\mathbf{y} \in V$.
3. It is as hard as the decision problem for the first-order theory over the word problem discussed below, which is known to be undecidable (Martin Davis, personal communication).
4. A point set is *closed regular* if it is equal to the closure of its interior. It is *open regular* if it is equal to the interior of its closure. The two categories are isomorphic, so it does not matter which is used.

Appendix: Proofs

Proof of Theorem 1

Definition 33 *Let* $\mathbf{T}$ be a topological space and let Ω be a collection of homogeneous partitions of $\mathbf{T}$. Let $\Theta = \bigcup_{\mathcal{Q}\in\Omega}\mathcal{Q}$; that is Θ is the set of all the cells that are in any partition in Ω. A sequence $\mathbf{U}_0 \ldots \mathbf{U}_k$ of cells in Θ is a chain *through* Ω if $\mathbf{U}_{i-1} \cap \mathbf{U}_i \neq \varnothing$ for $i = 1 \ldots k$. If $\mathbf{x} \in \mathbf{U}_0$ and $\mathbf{y} \in \mathbf{U}_k$, then we say that the chain connects $\mathbf{x}$ and $\mathbf{y}$.

Lemma 15 *Let* $\mathbf{x}$ be a point in $\mathbf{T}$, let $\mathcal{U}$ be a locally finite partition of $\mathbf{T}$, and let $\mathbf{V}$ be a subset of $\mathbf{T}$ such that $\mathbf{x} \in \mathrm{Cl}(\mathbf{V})$. Then there exists a cell $\mathbf{U} \in \mathcal{U}$ such that $\mathbf{U} \cap \mathbf{V} \neq \varnothing$ and $\mathbf{x} \in \mathrm{Cl}(\mathbf{U})$.

Proof: Since $\mathcal{U}$ is locally finite, let $\mathbf{N}$ be a neighborhood of $\mathbf{X}$ that intersects only finitely many cells in $\mathcal{U}$; let these be $\mathbf{U}_1 \ldots \mathbf{U}_k$. Since $\mathcal{U}$ is a partition of $\mathbf{T}$, it follows that $\mathbf{N} \subset \mathbf{U}_1 \cup \ldots \cup \mathbf{U}_k$; hence $\mathbf{N} \cap \mathbf{V} \subset (\mathbf{U}_1 \cap \mathbf{V}) \cup \ldots \cup (\mathbf{U}_k \cap \mathbf{V})$. Since the right hand is a finite union, $Cl(\mathbf{N} \cap \mathbf{V}) \subset Cl(\mathbf{U}_1 \cap \mathbf{V}) \cup \ldots \cup Cl(\mathbf{U}_k \cap \mathbf{V})$. Since $\mathbf{x} \in Cl(\mathbf{N} \cap \mathbf{V})$, it must be the case that $\mathbf{x} \in Cl(\mathbf{U}_i \cap \mathbf{V})$ for at least one of the $\mathbf{U}_i$. It is immediate that $\mathbf{x} \in Cl(\mathbf{U}_i)$ and $\mathbf{U}_i \cap \mathbf{V} \neq \varnothing$. ▪

Lemma 16 *Let* $\mathbf{T}$ be a topological space and let Ω be a non-empty collection of homogeneous partitions of $\mathbf{T}$. Let $\Theta = \bigcup_{\mathcal{Q}\in\Omega}\mathcal{Q}$. Define an equivalence relation over elements of $\mathbf{T}$, $\mathbf{x} \sim_\Omega \mathbf{y}$ if $\mathbf{x}$ and $\mathbf{y}$ are connected by a chain through Ω. It is immediate that this is an equivalence relation. Let $\mathcal{Q}$ be the collection of equivalence classes of $\mathbf{T}$ under $\sim_\Omega$. Then $\mathcal{Q}$ is a homogeneous partition of $\mathbf{T}$.

Proof: Since $\sim_\Omega$ is an equivalence relation, it is immediate that $\mathcal{Q}$ is a partition.

It is likewise immediate that, if $\mathbf{U}_0 \ldots \mathbf{U}_k$ is a chain through Ω, then there exists a cell $\mathbf{U} \in \mathcal{Q}$ that contains their union. In particular, each of the partitions in Ω is a refinement of $\mathcal{Q}$.

We need to prove that $\mathcal{Q}$ satisfies four properties.

First, $\mathcal{Q}$ is locally finite. Let $\mathbf{x}$ be a point in $\mathbf{T}$. Let $\mathcal{U}$ be a partition in Ω. Since $\mathcal{U}$ is locally finite, there is a neighborhood $\mathbf{N}$ of $\mathbf{x}$ that intersects only finitely many cells in $\mathcal{U}$. Since $\mathcal{U}$ is a refinement of $\mathcal{Q}$, it follows that $\mathbf{N}$ intersects only finitely many elements of $\mathcal{Q}$.

Second, $\mathcal{Q}$ allows simple transitions. Let $\mathbf{x}$ be any point in $\mathbf{T}$ and let $\mathbf{V}$ be a cell in $\mathcal{Q}$ such that $\mathbf{x} \in Cl(\mathbf{V})$. Let $\mathcal{U}$ be any partition in Ω. By lemma 15, there exists a cell $\mathbf{U} \in \mathcal{U}$ such that $\mathbf{x} \in Cl(\mathbf{U})$ and $\mathbf{U} \cap \mathbf{V} \neq \varnothing$. Since $\mathcal{U}$ is a refinement of $\mathcal{Q}$, we must have $\mathbf{U} \subset \mathbf{V}$. Since $\mathcal{U}$ is locally simple, there is a path π such that π has a starting transition from $\mathbf{x}$ to $\mathbf{U}$. But then π has a starting transition from $\mathbf{x}$ to $\mathbf{V}$. Thus $\mathcal{Q}$ allows simple transitions.

Third, $\mathcal{Q}$ is path-connected. Let $\mathbf{Q}$ be a cell in $\mathcal{Q}$ and let $\mathbf{x}$ and $\mathbf{y}$ be points in $\mathbf{Q}$. Then there is a chain $\mathbf{U}_0 \ldots \mathbf{U}_k$ through Ω connecting $\mathbf{x}$ and $\mathbf{y}$; as remarked above, all the $\mathbf{U}_i$ are subsets of $\mathbf{Q}$. For $i = 1 \ldots k-1$ let $\mathbf{z}_i$ be a point in $\mathbf{U}_i \cap \mathbf{U}_{i+1}$; let $\mathbf{z}_0 = \mathbf{x}$ and let $\mathbf{z}_k = \mathbf{x}$. Since $\mathbf{z}_i$ and $\mathbf{z}_{i+1}$ are in $\mathbf{U}_i$, which is path-connected, there is a path π_i from $\mathbf{z}_i$ to $\mathbf{z}_{i+1}$ that stays in $\mathbf{U}_i$ and hence in $\mathbf{Q}$. Splicing the π_i together gives a path from $\mathbf{x}$ to $\mathbf{y}$ that stays in $\mathbf{Q}$.

Fourth, $\mathcal{Q}$ is uniform in its neighbors. Let $\mathbf{Q}$ be a cell in $\mathcal{Q}$; let $\mathbf{x}$ and $\mathbf{y}$ be points in $\mathbf{Q}$, and let $\mathbf{V}$ be a cell in $\mathcal{Q}$ such that $\mathbf{x} \in Cl(\mathbf{V})$. We need to show that $\mathbf{y} \in Cl(\mathbf{V})$. Define the cells $\mathbf{U}_i$ and the points $\mathbf{z}_i$ as in the previous paragraph. Let $\mathcal{U}_i$ be the partition in Ω containing $\mathbf{U}_i$. By lemma 15, there is a cell $\mathbf{W}_1 \in \mathcal{U}_1$ such that $\mathbf{z}_0 \in Cl(\mathbf{W}_1)$, and $\mathbf{V} \cap \mathbf{W}_1 \neq \varnothing$. Since $\mathcal{U}_1$ is a refinement of $\mathcal{Q}$, $\mathbf{W}_1 \subset \mathbf{V}$. Again, there is a cell $\mathbf{W}_2 \in \mathcal{U}_2$ such that $\mathbf{z}_1 \in Cl(\mathbf{W}_2)$, and $\mathbf{W}_1 \cap \mathbf{W}_2 \neq \varnothing$; and again $\mathbf{W}_2 \subset \mathbf{V}$. Continuing on in this way, we can construct a chain $\mathbf{W}_1, \mathbf{W}_2 \ldots \mathbf{W}_k$ through Ω such that $\mathbf{W}_i \subset \mathbf{V}$ and $\mathbf{z}_i$ and $\mathbf{z}_{i+1}$ are in $Cl(\mathbf{W}_i)$. Thus $\mathbf{y} \in Cl(\mathbf{V})$. ▪

Theorem 1 *Let* $\mathcal{U}$ be a partition of $\mathbf{T}$. If there exists a homogeneous refinement of $\mathcal{U}$, then there exists a QHD of $\mathcal{U}$.

Proof: Let Ω be the collection of all homogeneous refinements of $\mathcal{U}$, and construct $\mathcal{Q}$ as in lemma 16. By lemma 16, $\mathcal{Q}$ is a homogeneous partition of $\mathbf{T}$. It is immediate that $\mathcal{Q}$ is a refinement of $\mathcal{U}$ and that every homogeneous refinement of $\mathcal{U}$ is a refinement of $\mathcal{Q}$.

Proof of Theorem 2

Lemma 17 *Let* $\mathcal{U}$ be a partition of $\mathbf{T}$, and let $\mathcal{V}$ be a homogeneous refinement of $\mathcal{U}$. Then $\mathcal{V}$ is a refinement of both $\Phi(\mathcal{U})$ and $\Psi(\mathcal{U})$.

Proof: First, let $\mathbf{V}$ be a cell of $\mathcal{V}$ and let $\mathbf{x}$ and $\mathbf{y}$ be points in $\mathbf{V}$. Since $\mathbf{V}$ is path-connected, there is a path π from $\mathbf{x}$ to $\mathbf{y}$ that remains in $\mathbf{V}$. Since $\mathcal{V}$ is a refinement of $\mathcal{U}$, there exists a cell $\mathbf{U} \in \mathcal{U}$ such that $\mathbf{V} \subset \mathbf{U}$. Then π remains in $\mathbf{U}$; therefore $\mathbf{V}$ is a subset of one path-connected component of $\mathbf{U}$. Therefore $\mathcal{V}$ is a refinement of $\Phi(\mathcal{U})$.

Second, let $\mathbf{V}$, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{U}$ be as above. Let $\mathbf{W}$ be a cell of $\mathcal{U}$ such that $\mathbf{x} \in Cl(\mathbf{W})$; we wish to show that $\mathbf{y} \in Cl(\mathbf{W})$. Since $\mathcal{V}$ is a refinement of $\mathcal{U}$ and since both $\mathcal{V}$ and $\mathcal{U}$ are locally finite, there exists a cell $\mathbf{P} \in \mathcal{V}$ such that $\mathbf{P} \subset \mathbf{W}$ and $\mathbf{x} \in Cl(\mathbf{P})$. Since $\mathcal{V}$ is uniform in neighbors, $\mathbf{y} \in Cl(\mathbf{P})$; thus $\mathbf{y} \in Cl(\mathbf{W})$ as desired. Thus, by symmetry $\mathbf{N}(\mathbf{x}, \mathcal{U}) = \mathbf{N}(\mathbf{y}, \mathcal{U})$, so $\mathbf{x}$ and $\mathbf{y}$ are in the same cell of $\Psi(\mathcal{U})$. Hence $\mathcal{V}$ is a refinement of $\Psi(\mathcal{U})$.

Theorem 2 *Let* $\mathcal{U}$ be a locally simple partition of $\mathbf{T}$. Let $\mathcal{U}_\sigma$ be the decompositional sequence of $\mathcal{U}$. Then:

a. *The sequence reaches a fixed point. That is, there exists an ordinal* τ such that, for all $\sigma > \tau$, $\mathcal{U}_\sigma = \mathcal{U}_\tau$.
b. *If* $\mathcal{U}_\tau$ is locally simple, then it is the QHD of $\mathcal{U}$.
c. *If there exists a QHD of* $\mathcal{U}$, then it is $\mathcal{U}_\tau$.

Proof: First, we prove by transfinite induction that, if $\sigma > \nu$ then $\mathcal{U}_\sigma$ is a refinement of $\mathcal{U}_\nu$. Assume that the statement holds for all $\eta < \sigma$. If σ is the successor of μ, and $\nu < \sigma$ then $\nu \leq \mu$ so $\mathcal{U}_\mu$ is a refinement of $\mathcal{U}_\nu$ by the inductive hypothesis. Since $\Phi(\mathcal{V})$ and $\Psi(\mathcal{V})$ are refinements of $\mathcal{V}$ for any $\mathcal{V}$, it follows that $\mathcal{U}_\sigma = \Psi(\Phi(\mathcal{U}_\mu))$ is a refinement of $\mathcal{U}_\mu$ and hence of $\mathcal{U}_\mu$.

If σ is a limit ordinal, then the conclusion is immediate from the definition.

Second, since each partition of $\mathbf{T}$ is a set of subsets of $\mathbf{T}$, the cardinality of the set of partitions of $\mathbf{T}$ is certainly no more than $2^{2^{|\mathbf{T}|}}$. Since refinement is a partial ordering over partitions, a chain of refinements cannot have cycles, and consequently cannot have more than more than $2^{2^{|\mathbf{T}|}}$ different values. (It can easily be shown that the true bound is actually $|\,\mathbf{T}\,|$, but we do not need that here.) Hence a fixed point must be reached for ordinals corresponding to that cardinality. This establishes part (a).

Third, since $\mathcal{U}_\tau$ is a fixed point of $\Psi(\Phi(\cdot))$, if it is also locally simple, then it is homogeneous, by definition. By a simple transfinite induction using lemma 17, it follows that, if $\mathcal{V}$ is any homogeneous refinement of $\mathcal{U}$, then it is a refinement of $\mathcal{U}_\tau$. Hence, if $\mathcal{U}$ is locally simple, then it is the QHD of $\mathcal{U}$. This establishes (b). Conversely, if there exists a QHD $\mathcal{W}$ of $\mathcal{U}$, then it is a refinement of $\mathcal{U}_\tau$. Thus $\mathcal{U}_\tau$ must be locally simple; hence it is a homogeneous refinement of $\mathcal{U}$; hence it must be equal to $\mathcal{W}$. This establishes (c).

Proof of Theorems 13 and 14

The construction in this proof is a little involved, so we will begin with a general overview and some motivation. Throughout this section, $\mathcal{U}$ is a path-connected partition of $\mathbf{T}$ and G is the corresponding transition graph.

The proof proceeds in two steps. First, we slightly extend the domain of transition graphs and gpaths, and we define a structure over this extension that is first-order equivalent to the structure $\prec \text{DPATH}_U, \mathcal{L}^p, \mathcal{I}_U^p \succ$ over domain of paths in $\mathbf{T}$. Second, we show that this extension can be modelled in the unextended structure $\prec \text{DGRAPH}_G, \mathcal{L}^g, \mathcal{I}_G^g \succ$ over the domain of gpaths through the transition graph.

The obvious difference between the DPATH domain and the DGRAPH domain is that the former has points and the latter does not. So we will need to extend the latter to have something corresponding to points; we will call these "gpoints."

Moreover, if a cell has infinitely many points, then, in a first-order language with equality, it can be stated that it has at least k points for any finite k, and this must be true of the gpoints as well. Therefore, for any cell with infinitely many points, there must exist infinitely many gpoints. For a cell with a single point, such as cell **B1a1** of Table 2, there must exists a single gpoint. Therefore, we adopt the following definition:

Definition 34 *Let* Ω be an arbitrary infinite set. A gpoint *of* $\mathbf{U}$ is any pair $\langle \mathbf{U}, \omega \rangle$ where $\omega \in \Omega$. If $\mathbf{U}$ is an infinite cell, then every such pair is a gpoint of $\mathbf{U}$. If $\mathbf{U}$ is a singleton cell, then there is one such pair that is a gpoint of $\mathbf{U}$. The set of gpoints of $\mathbf{U}$ is denoted "GPS($\mathbf{U}$)." If $\mathbf{x}$ is a gpoint of $\mathbf{U}$, then we say that $\mathbf{U}$ is the owner of $\mathbf{x}$, written $\mathbf{U} = \mathbf{O}(\mathbf{x})$.

Typographically, we will use boldface, lower case symbols for gpoints as well as points.

We also need to construct extended gpaths, called "egpaths" to deal with two issues. First, a path starts and ends in a point; so a egpath must start and end in a gpoint. Second, there is again an issue with the number of paths. For any starting point $\mathbf{x} \in \mathbf{U}$, ending point $\mathbf{y} \in \mathbf{V}$ and gpath β from $\mathbf{U}$ to $\mathbf{V}$, there normally exist infinitely many paths π that start at $\mathbf{x}$, end at $\mathbf{y}$, and have trace β; so there must exist infinitely many egpaths of this kind as well. There is one exception: if $\mathbf{U}$ is a singleton cell, then there is a single path that stays in $\mathbf{U}$.

Definition 35 *A* singleton path *is a path that remains in a single singleton cell.*

Let Ω and ω_0 be as above. Let β be a gpath. An *egpath corresponding to* β is a quadruple $\langle \beta, \mathbf{x}, \mathbf{y}, \omega \rangle$ where $\mathbf{x} \in GPS(S(\beta))$, $\mathbf{y} \in GPS(E(\beta))$, and $\omega \in \Omega$. If β is not a singleton path, then any such quadruple is a egpath for β. If β is a singleton path, then there is one such quadruple that is the unique egpath for β. The set of all egpaths corresponding to β is written "EGPATHS(β)." If $\gamma \in$ EGPATHS(β) we will write $\beta = \mathbf{O}(\gamma)$.

The only properties of gpoints and egpaths that matter are (a) that a gpoint is in a cell; (b) that a gpoint is the start or end of an egpath; (c) that there are the right number (one or infinite) of gpoints and egpaths for each cell or gpath respectively.

We now define a collection of relations on the space of gpoints, cells, and egpaths.

EGPOINTS_G = the union of GPS($\mathbf{V}$) for $\mathbf{V} \in CELLS_G$.

EGPATHS_G = the union of EGPATHS(β) for $\beta \in \text{GPATHS}_G$

$\text{DEGPATHS}_G = \text{GPOINTS}_G \cup \text{CELLS}_G \cup \text{EGPATHS}_G$.

$\text{EGIN}_G = \{\prec \mathbf{x}, \mathbf{V} \succ \mid \mathbf{x} \in \text{GPS}(\ \mathbf{V}\ )\ \}$.

$\text{EGSTART}_G = \{\prec \gamma, \mathbf{V} \succ \mid \mathbf{O}(\gamma) \in \text{GPATHS}_G \wedge \mathbf{V} = S(\mathbf{O}(\gamma))\,\}$.

$\text{EGEND}_G = \{\prec \gamma, \mathbf{V} \succ \mid \mathbf{O}(\gamma) \in \text{GPATHS}_G \wedge \mathbf{V} = S(\mathbf{O}(\gamma))\,\}$.

$\text{EGREMAINS}_G = \{\prec \gamma, \mathbf{V} \succ \mid \mathbf{O}(\gamma) = \langle \mathbf{V} \rangle)\ \}$

$\text{EGCLOSEDTRANS}_G = \{\prec \gamma, \mathbf{V}, \mathbf{W} \succ \mid \prec \mathbf{O}(\gamma), \mathbf{V}, \mathbf{W} \succ \in \text{FORWARDARC}_G\}$

$\text{EGOPENTRANS}_G = \{\prec \gamma, \mathbf{V}, \mathbf{W} \succ \mid \prec \mathbf{O}(\gamma), \mathbf{V}, \mathbf{W} \succ \in \text{BACKWARDARC}_G\}$

$\text{EGSPLICE}_G = \{\prec \gamma_1, \gamma_2, \gamma_3 \succ \mid \mathbf{O}(\gamma_1) \mid \mathbf{O}(\gamma_2) = \mathbf{O}(\gamma_3)\}$.

Define the interpretation $\mathcal{I}^e$ of $\mathcal{L}^p$ as mapping each symbol onto the corresponding relation over $DEGPATH_U$ and as mapping the constant symbols U_i onto the same cells as $\mathcal{I}^p$. Let $\mathcal{S}^p$ be the structure $\prec \text{DPATHS}_U, \mathcal{L}^p, \mathcal{I}^p_U \succ$ and let $\mathcal{S}^e$ be the structure $\prec \text{DEGPATHS}_G, \mathcal{L}^p, \mathcal{I}^e_G \succ$

Definition 36 *A subset* MP of $DPATHS_U$ is matchable *if it satisfies the following:*

- *MP is the union of* $CELLS_U$ *with a finite (possibly empty) set of points and a finite set of paths.*
- *If path* $\pi \in MP$ *, then* $S(\pi) \in MP$ *and* $E(\pi) \in \text{MP}$.

A subset ME of $DEGPATHS_G$ *is* matchable *if it satisfies the following:*

- *MP is the union of* $CELLS_G$ *with a finite (possibly empty) set of gpoints and a finite set of egpaths.*
- *If egpath* $\gamma = \langle \beta, \mathbf{x}, \mathbf{y}, \omega \rangle \in ME$ *then* $\mathbf{x} \in ME$ *and* $\mathbf{y} \in ME$.

Definition 37 *Let* MP be a matchable subset of $DPATHS_U$ and let ME be a matchable subset of $DEGPATHS_U$. A bijection ζ from MP to ME is a correspondence *if for all* $m, n \in MP$:

- If $m \in CELLS_U$ then $\zeta(m) = m$.
- If $m \in POINTS_U$ then $\zeta(m) \in EGPOINTS_G$ and $\mathbf{O}(m, \mathcal{U}) = \mathbf{O}(\zeta(m))$.
- If $m \in PATHS_U$ then $\zeta(m) \in EGPATHS_G$ and $\mathbf{O}(\zeta(m)) = \Delta(\Gamma_{U,Z4}(m))$.
- If $m \in PATHS_U$ and $n \in POINTS_U$ then

[$n = S(m)$ if and only if $\prec \zeta(m), \zeta(n) \succ \in EGSTART_G$] and
[$n = E(m)$ if and only if $\prec \zeta(m), \zeta(n) \succ \in EGEND_G$].

Definition 38 *Let* $\mu_1 \ldots \mu_k$ be variable symbols. Let σ be a valuation of the μ_i in $DPATHS_U$ and let τ be a valuation of the μ_i in $DEGPATHS_G$. We say that σ and τ correspond *if there exist set* MP and ME and a correspondence ζ such that for each variable μ_i, $\sigma(\mu_i) \in MP$, $\tau(\mu_i) \in ME$, and $\tau(\mu_i) = \zeta(\sigma(\mu_i))$

From here the proof proceeds along standard lines for proof of first-order equivalence. We show that any matching valuations have extensions that are still matching; we show that atomic formulas are equivalent under matching valuations; and we show inductively that complex formulas are equivalent under matching valuations.

Lemma 18 *Let* σ and τ be corresponding valuations, and let ϕ be an atomic formula in $\mathcal{I}^p$. Then $\mathcal{S}_U^p, \sigma \models \phi$ if and only if $\mathcal{S}_U^e, \tau \models \phi$.

Proof: Straightforward from the definitions, though lengthy. Check each predicate and equality in turn.

Corollary 19 *Let* σ and τ be corresponding valuations, and let ϕ be a quantifier-free formula in $\mathcal{I}^p$. Then $\mathcal{S}_U^p, \sigma \models \phi$ if and only if $\mathcal{S}_U^e, \tau \models \phi$.

Proof: Immediate from lemma 18.

Definition 39 *If* σ is a valuation and μ is a variable not in the domain of σ, then an extension of σ to μ *is a valuation that agrees with* σ on all the variables in the domain of σ and assigns a domain value to μ.

Lemma 20 *Let* σ and τ be corresponding valuations. Let σ' be an extension of σ to a new variable μ. Then there exists an extension τ' of τ to μ such that τ' corresponds to σ'.

Conversely, if τ' *is an extension of* τ *to* μ *then there exists an extension* σ' to μ that corresponds to τ.

Proof of the first implication: Let ζ be a correspondence matching σ and τ and let MP and ME be the associated matchable sets. If $\sigma'(\mu) \in MP$, then define $\tau'(\mu) = \zeta(\sigma'(\mu))$.

Otherwise, let $m = \sigma'(\mu)$.

If m is a point, define $\zeta'(m)$ to be $\zeta(m)$ if $m \in MP$; otherwise, to be a gpoint n such that $\mathbf{O}(n) = \mathbf{O}(m, \mathcal{U})$ and $n \not\in ME$. Let τ' be the extension of τ with $\tau'(\mu) = n$.

If m is a path, define $\zeta'(m)$ to be $\zeta(m)$ if $m \in MP$; otherwise, to be a egpath n such that $\mathbf{O}(n) = \Delta(\Gamma_{U,Z4}(m))$ and $n \not\in ME$. Define $\zeta'(S(m))$ and $\zeta'(E(m))$ as above. Let $\tau'(\mu) = n$.

It is immediate that ζ' is a correspondence and that σ' and τ' are corresponding valuations. Note that this construction relies, first, on the fact that there are infinitely many gpoints for every infinite cell and infinitely many egpaths for every gpath and every starting and ending gpoint; and, second, that it relies on theorem 4 that there exists a gpath corresponding to every path.

The proof of the second implication has exactly the same structure. Here, in going from egpaths to paths, we rely on corollary 7 to be sure that there is a path corresponding to every gpath and starting and ending point. ▪

Lemma 21 *Let* σ and τ be corresponding valuations, and let ϕ be a formula in $\mathcal{I}^p$. Then $\mathcal{S}^p_U, \sigma \vDash \phi$ if and only if $\mathcal{S}^e_G, \tau \vDash \phi$.

Proof by induction on the number k of quantifiers in ϕ. Assume that ϕ has been placed in prenex form. The case $k = 0$ is just corollary 19.

Suppose that ϕ has the form $\exists_\mu \psi$. Let free variables $\mu_1 \dots \mu_q$ be the free variables of ϕ. If $\mathcal{S}^p_U, \sigma \vDash \phi$ then let m be an entity in DPATHS_U such that the extension of σ, $\sigma' = \sigma \cup \{\mu \rightarrow m\}$ satisfies $\mathcal{S}^p_U, \sigma' \vDash \psi$. By lemma 20 there exists a valuation τ' extending τ corresponding to σ'. By the inductive hypothesis, $\mathcal{S}^e_G, \tau' \vDash \psi$. Therefore $\mathcal{S}^e_G, \tau \vDash \phi$.

The converse --- if $\mathcal{S}^e_G, \tau \vDash \phi$ then $\mathcal{S}^p_U, \sigma \vDash \phi$ --- is exactly analogous.

Suppose that ϕ has the form $\forall_\mu \psi$. Then $\mathcal{S}^p_U, \sigma \vDash \phi$ if and only if $\mathcal{S}^p_U, \sigma \vDash \exists_\mu \neg\psi$ which by the contrapositive to the previous paragraph holds if and only if $\mathcal{S}^e_G, \tau \vDash \exists_\mu \neg\psi$, which holds if and only if $\mathcal{S}^e_G, \tau \vDash \phi$.

Corollary 22 *The structures* $\mathcal{S}^p_U$ and $\mathcal{S}^e_G$ are first-order equivalent; i.e, for any sentence $\phi \in \mathcal{L}^p$, $\mathcal{S}^p_U \vDash \phi$ if and only if $\mathcal{S}^e_G \vDash \phi$.

Proof: This is just the special case of lemma 21 for formulas with no free variables. ▪

We now show that the structure $\mathcal{S}^e_G$ is definable in $\mathcal{S}^g_G$, using a definitional mapping that is independent of G. To do this, we will have to model gpoints and egpaths in terms of cells and gpaths, which are all we have in $\mathcal{S}^g_G$, and we have to translate the predicates (including equality and the universal relation) of $\mathcal{L}^p$ into corresponding formulas in $\mathcal{L}^g$. Neither of these is very difficult.

First, we need to model Ω. Since all we need out of Ω is that it should be infinite and that we can determine equality and inequality, we can just use GPATHS_G itself, since there are always infinitely many gpaths (This is why we require that $\mathcal{U}$ has at least two cells).

If $\mathbf{x} = \langle \mathbf{U}, \omega \rangle$ is a singleton gpoint—i.e. the unique gpoint in singleton cell $\mathbf{U}$—then we choose ω to be the gpath $\langle \mathbf{U} \rangle$. If $\gamma = \langle \beta, \mathbf{x}, \mathbf{x}, \omega \rangle$ is a singleton path—i.e the unique path that remains in a singleton cell $\mathbf{U}$—then again then we choose ω to be the gpath $\langle \mathbf{U} \rangle$.

We map any entity $m \in \text{DEGPATHS}_G$ to a quadruple over DGPATHS_G.

Definition 40 *Define the mapping* $\Theta : DEGPATHS_G \rightarrow (DGPATHS_G)^4$ as follows:

- If m is the cell $\mathbf{U}$ then $\Theta(m) = \prec \mathbf{U}, \mathbf{U}, \mathbf{U}, \mathbf{U} \succ$.
- If m is the gpoint $\langle \mathbf{U}, \omega \rangle$ then $\Theta(m) = \prec \mathbf{U}, \omega, \mathbf{U}, \mathbf{U} \succ$.
- If m is the egpath $\langle \beta, \mathbf{x}, \mathbf{y}, \omega_1 \rangle$ where $\mathbf{x} = \langle S(\beta), \omega_2 \rangle$ and $\mathbf{y} = \langle E(\beta), \omega_3 \rangle$

then $\Theta(m) = \prec \beta, \omega_1, \omega_2, \omega_3 \succ$.

Note that if m is a singleton gpoint, then m_1 is a singleton cell, Remains(m_2, m_1) and $m_1 = m_3 = m_4$. If m is a singleton egpath, then there exists a singleton cell $\mathbf{U}$ such that Remains($m_1, \mathbf{U}$), and $m_1 = m_2 = m_3 = m_4$.

The following rules specify the translation of each relation of $\mathcal{S}_G^e$ into a first-order definable relation in $(\mathcal{S}_G^g)^4$. We will write $\vec{m}$ as an abbreviation for the tuple $\prec m_1, m_2, m_3, m_4 \succ$. For readability, we will omit the subscript G, which applies to all the relations below.

$$\Theta(\mathbf{U}_i) = \prec \mathbf{U}_i, \mathbf{U}_i, \mathbf{U}_i, \mathbf{U}_i \succ$$

$$\Theta(\text{CELLS}) = \{\vec{m} \,|\, m_1 \in \text{CELLS} \wedge m_1 = m_2 = m_3 = m_4\}$$

$$\Theta(\text{SINGLETON}) = \{\vec{m} \,|\, \vec{m} \in \Theta(\text{CELLS}) \wedge m_1 \in \text{SINGLETON})\}$$

$$\begin{aligned} &\Theta(\text{GPOINTS}) = \\ &\{\vec{m} \,|\, [m_1 \in \text{SINGLETON} \wedge \prec m_1, m_2 \succ \in \text{REMAINS} \wedge m_1 = m_3 = m_4] \vee \\ &\qquad [m_1 \in \text{CELLS} - \text{SINGLETON} \wedge m_2 \in \text{GPATHS} \wedge m_1 = m_3 = m_4] \\ &\} \end{aligned}$$

The definition of Θ (EGPATHS) is rather complicated, because we have to deal, both with singleton egpaths and with the cases of non-singleton egpaths that begin or end at a singleton point. Moveover, we have structured this definition to be purely existential, so as to support a translation of an existential formula in $\mathcal{L}^p$ into an existential formula in $\mathcal{L}^g$; the translation could be somewhat simpler if we allowed the use of universal quantifiers. We begin by defining some relations over DGPATHS.

SINGLETONPATH = $\{\gamma \,|\exists_{\text{U}} \mathbf{U} \in \text{SINGLETON} \wedge \prec \gamma, \mathbf{U} \succ \in \text{GREMAINS}\}$

NONSINGLETONPATH =

$\{\gamma \,|\, \exists_{\text{U}} \prec \gamma, \mathbf{U} \succ \in \text{GSTART} \wedge [\mathbf{U} \not\in \text{SINGLETON} \vee \prec \gamma, \mathbf{U} \succ \not\in \text{GREMAINS}]\}$

SINGLETONSTART = $\{\prec \gamma, \mathbf{U} \succ | \prec \gamma, \mathbf{U} \succ \in \text{GSTART} \wedge \mathbf{U} \in \text{SINGLETON}\}$

NONSINGLETONSTART = $\{\gamma \,|\exists_{\text{U}} \prec \gamma, \mathbf{U} \succ \in \text{GSTART} \wedge \mathbf{U} \not\in \text{SINGLETON}\}$

SINGLETONEND = $\{\prec \gamma, \mathbf{U} \succ | \prec \gamma, \mathbf{U} \succ \in \text{GEND} \wedge \mathbf{U} \in \text{SINGLETON}\}$

NONSINGLETONEND = $\{\gamma \,|\exists_{\text{U}} \prec \gamma, \mathbf{U} \succ \in \text{GEND} \wedge \mathbf{U} \not\in \text{SINGLETON}\}$

$$
\begin{aligned}
&\Theta(\text{EGPATHS}) = \\
&\{\vec{m} \mid [m_1 \in \text{SINGLETONPATH} \wedge m_1 = m_2 = m_3 = m_4] \vee \\
&\quad [m_1, m_2, m_3, m_4 \in \text{GPATHS} \wedge m_1 \in \text{NONSINGLETONPATH} \wedge \\
&\quad\; [m_1 \in \text{NONSINGLETONSTART} \vee \\
&\quad\;\; [\exists_{\text{U}} \prec m_1, \mathbf{U} \succ \in \text{SINGLETONSTART} \wedge \prec m_3, \mathbf{U} \succ \in \text{GREMAINS}]] \wedge \\
&\quad\; [m_1 \in \text{NONSINGLETONEND} \vee \\
&\quad\;\; [\exists_{\text{U}} \prec m_1, \mathbf{U} \succ \in \text{SINGLETONEND} \wedge \prec m_4, \mathbf{U} \succ \in \text{GREMAINS}]] \\
&\quad] \\
&\}
\end{aligned}
$$

$\Theta(\text{DEGPATHS}) = \Theta(\text{CELLS}) \cup \Theta(\text{GPOINTS}) \cup \Theta(\text{EGPATHS})$.

Θ (EGIN) = $\{ \prec \vec{m}, \vec{n} \succ \mid \vec{m} \in \Theta\,(\text{GPOINTS}) \wedge \vec{n} \in \Theta\,(\text{CELLS}) \wedge m_1 = n_1 \}$.

Θ (EGSTART) = $\{ \prec \vec{m}, \vec{n} \succ \mid \vec{m} \in \Theta\,(\text{EGPATHS}) \wedge \vec{n} \in \Theta\,(\text{GPOINTS}) \wedge \prec m_1, n_1 \succ \in \text{GSTART} \}$.

Θ (EGEND) = $\{ \prec \vec{m}, \vec{n} \succ \mid \vec{m} \in \Theta\,(\text{EGPATHS}) \wedge \vec{n} \in \Theta\,(\text{GPOINTS}) \wedge \prec m_1, n_1 \succ \in \text{GEND} \}$.

Θ (EGREMAINS) =

$\{ \prec \vec{m}, \vec{n} \succ \mid \vec{m} \in \Theta\,(\text{EGPATHS}) \wedge \vec{n} \in \Theta\,(\text{GPOINTS}) \wedge \prec m_1, n_1 \succ \in \text{GREMAINS} \}$.

Θ (EGCLOSEDTRANS) =

$\{ \prec \vec{m}, \vec{n}, \vec{p} \succ \mid \vec{m} \in \Theta\,(\text{EGPATHS}) \wedge \vec{n}, \vec{p} \in \Theta\,(\text{CELLS}) \wedge \prec m_1, n_1, p_1 \succ \in \text{FORWARDARC} \}$.

Θ (EGOPENTRANS) =

$\{ \prec \vec{m}, \vec{n}, \vec{p} \succ \mid \vec{m} \in \Theta\,(\text{EGPATHS}) \wedge \vec{n}, \vec{p} \in \Theta\,(\text{CELLS}) \wedge \prec m_1, n_1, p_1 \succ \in \text{BACKWARDARC} \}$.

Θ (EGSPLICE) = $\{ \prec \vec{m}, \vec{n}, \vec{p} \succ \mid \vec{m}, \vec{n}, \vec{p} \in \Theta\,(\text{EGPATHS}) \wedge \prec m_1, n_1, p_1 \succ \in \text{GSPLICE} \}$.

Θ (=) = $\{ \prec \vec{m}, \vec{n} \succ \mid \vec{m}, \vec{n} \in \Theta\,(\text{DEGPATHS}) \wedge \vec{m} = \vec{n} \}$.

That completes the definition of Θ.
Lemma 23 *For any relation* Φ in $\mathcal{S}_G^e$ and entities $x_1 \ldots x_k \in DEGPATHS_G$,
$\prec x_1 \ldots x_k \succ \in \Phi$ if and only if $\prec \Theta(x_1) \ldots \Theta(x_k) \succ \in \Theta(\Phi)$.
Proof: Long but straightforward case analysis. Each case follows immediately from the definition.

Lemma 24 *The structure* $\mathcal{S}^e_G$ is first-order definable in terms of $\mathcal{S}^g_G$. Moreover, the form of the definition is independent of G.

Proof: Immedate from lemma 23.

Theorem 10 *There exists a linear-time function* $\mathcal{A}^p$ that maps every sentence in $\mathcal{L}^p$ to a sentence in $\mathcal{L}^g$ satisfying the following. Let **T** be a Hausdorff space, let $\mathcal{U}$ be a path-connected partition over **T** with at least 2 cells, and let $\mathcal{I}^p_U$ be the interpretation of $\mathcal{L}^p$ in $DPATHS_U$ defined above. Let G be the transition graph corresponding to $\mathcal{U}$ and let $\mathcal{I}^g_G$ be the interpretation of $\mathcal{L}^g$ in $DGRAPH_G$ defined above, such that for each symbol U_i, $\mathcal{I}^g_G(U_i) = \mathcal{I}^p_U(U_i)$. Let Φ be any sentence in $\mathcal{L}^p$. Then Φ holds in the structure $\prec DPATH_U, \mathcal{L}^p, \mathcal{I}^p_U \succ$ if and only if $\mathcal{A}^p(\Phi)$ holds in the structure $\prec DGRAPH_G, \mathcal{L}^g, \mathcal{I}^g_G \succ$.

Proof: Immediate from corollary 22 and lemma 24. ▪

The proof of theorem 14 is exactly analogous. Extended ILS's are defined analogously to extended gpaths. One can then prove that the structure over extended ILS's is elementary equivalent to the Z5 structure of paths, and is definable in the structure of unextended ILS's.