REFERENCES

Bartlett, F. C. *Remembering*. Cambridge, Eng.: Cambridge University Press, 1932.

Bresnan, J. A realistic transformational grammar. In M. Halle, J. Bresnan & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, Mass.: MIT Press, 1978.

Card, S. K., Moran, T. P., & Newell, A. Computer text-editing: An information processing analysis of a routine cognitive skill. *Cognitive Psychology*, 1980, *12*, 32–74.

Card, S. K., Moran, T. P., & Newell, A. *The Psychology of Human-Computer Interaction*. Hillsdale, N.J.: Lawrence Erlbaum Assoc., in press.

Carroll, J. M., & Thomas, J. *Metaphor and the cognitive representation of computing systems*. Yorktown Heights, N.Y.: IBM Research Report RC-8302, 1980.

du Boulay, J. B. H., O'Shea, T., & Monk, J. The black box inside the glass box: Presenting computing concepts to novices. *International Journal of Man-Machine Studies*, 1981, *14*, 237–249.

Mayer, R. E. The psychology of how novices learn computer programming. *Computing Surveys*, 1981, *13*, 121–141.

Moran, T. P. An applied psychology of the user. *Computing Surveys*, 1981, *13*, 1–11. (a)

Moran, T. P. The Command Language Grammar: A representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, 1981, *15*, 3–50. (b)

Newell, A. Personal communication, 1980.

Newell, A., & Simon, H. A. *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Reisner, P. Using a formal grammar in human factors design of an interactive graphics system. *IEEE Transactions on Software Engineering, SE-7*, 1981, 229–240.

Self, J. Student models in CAI. *International Journal of Man-Machine Studies*, 1974, *6*, 261–276.

Sleeman, D. H., & Brown, J. S. Editorial: Intelligent tutoring systems. *International Journal of Man-Machine Studies*, 1979, *11*, 1–3.

Young, R. M. The machine inside the machine: Users' models of pocket calculators. *International Journal of Man-Machine Studies*, 1981, *15*, 51–85.

4 Qualitative Reasoning About Space and Motion

Kenneth D. Forbus
Massachusetts Institute of Technology

1. INTRODUCTION

People reason fluently about motion through space. For example, if two balls are thrown in to a well they might collide, but if one ball is always outside and the other always inside they cannot. The models we use in this qualitative kind of reasoning seem to be simpler than formal mechanics and appear to be based on our experience in the physical world.

One way to test theories about how we reason about a class of situations is to build a program that can answer the same sorts of questions we might ask about those situations. Writing a program requires explicit consideration of how the knowledge involved is used as well as just what must be known. If it works, the program provides a strong argument that the theory it embodies is sufficient for the domain. The behavior of the program can be compared to human performance on the questions of interest to see how well the theory it embodies explains what people do. Some of the details necessary to make a program run may suggest more precise experiments on human subjects.

There are several limitations to this methodology. Programs that do any reasoning at all are usually at the limits of current technology. Not all of the details of what the program does are relevant to what people do—they are a consequence of the different types of hardware available for the task (see (Marr, 1976). Worse yet, Computer Science is very young—much like chemistry before the periodic table, and certainly before biochemistry. Nevertheless, the ways of thinking about processes that it provides are the most precise we have.

The focus of this work on formalizing common sense knowledge is much in the spirit of the Naive Physics effort of Hayes (Hayes 1979a). However, Hayes ignores computational issues such as the use of a diagram and the "style" of reasoning, which are considered here. My approach is very different from the efforts described in (Bundy, 1976; Novak, 1976, and McDermott & Larkin, 1978), which are mainly concerned with modeling students solving textbook physics problems. When learning physics, students are forced to relate the new information to the physical knowledge they already have. It is only the latter kind of knowledge that is examined here.

1.1 The Domain and The Program

To explore the issues involved in reasoning about motion a program called FROB (Forbus, 1981a) was written. FROB reasons about motion in a simplified domain called the "Bouncing Ball" world. A situation in the Bouncing Ball world consists of a two dimensional scene with surfaces represented by line segments, and one or more balls which are modeled as point masses. A typical situation is depicted in Fig. 4.1. These assumptions allow us to avoid dealing with complex shapes and the third dimension. Only motion through space and momentary collisions with surfaces are considered; sliding, rolling, spinning, and other types of motion are ignored. Gravity is the sole external influence considered, for we wish to ignore air resistance and such complexities as charged or magnetic balls.

A scene is specified by a diagram containing a description of the surfaces. Given a scene, FROB analyzes the surface geometry and computes qualitative descriptions of the free space in the diagram. The person using the program can describe balls, properties of their states of motion, request simulations, and make global assumptions about the motion. FROB incrementally creates and updates

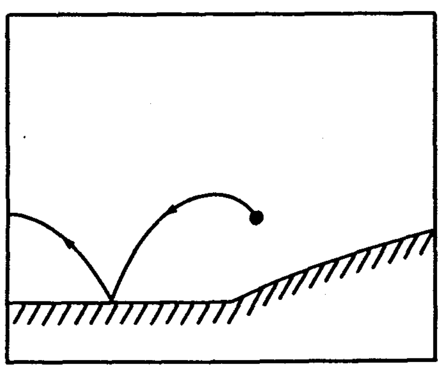


FIG. 4.1. *A typical scene from the Bouncing Ball world.* A situation in the Bouncing Ball World consists of a diagram that specifies surfaces and one or more balls. This drawing only shows the geometric aspects of the descriptions involved.

its descriptions to accommodate this information, complaining if inconsistencies are detected. Questions may be asked by calling procedures that interrogate these descriptions.

1. What can it (a ball) do next?
2. Where can it go next?
3. Where can it end up?
4. Can these two balls collide?

1.2 Main Ideas about Motion and Space

There are several theories concerning reasoning about motion and space that FROB illustrates. To summarize:

1. A quantitative "analog" geometric representation simplifies reasoning about space. It does so by providing a simple method for answering a class of geometric questions. Qualitative spatial reasoning can be thought of as manipulating a set of symbolic descriptions of space, defined in terms of the underlying analog representation.
2. Describing the motion of an object can be viewed as creating a network from descriptions of qualitatively distinct types of motion. They are linked by descriptions of the state of the object before and after each of these motions. This network can be used to analyze the motion and in some cases can be constructed by a process of simulation.
3. The result of envisioning (de Kleer, 1975, 1979) can be used as a device to assimilate assumptions about global properties of motion and in checking the actual motion of an object against these assumptions. The assimilation process makes heavy use of qualitative spatial descriptions and basic properties of motion.

2. SPATIAL DESCRIPTIONS

We do not yet know why people are so good at reasoning about space. Theorem proving and symbolic manipulation of algebraic expressions do not seem to account for this ability. Arguments against theorem proving may be found in [Waltz & Boggess, 1979], while the sheer complexity of algebraic manipulations argues against it as a basis for our spatial abilities. I conjecture that people find diagrams useful because they allow certain spatial questions to be decided by interpreting the results of perception. The marks in a diagram reflect the spatial relations between the things they represent, which allows us to use our visual apparatus to interpret these relationships as we would with real objects. In this case, perception provides a simple (at least for the processes that use it) decision procedure for a class of spatial questions.

People also can reason about space using less detailed representations than that of a diagram, as the well example discussed in the beginning of the paper illustrates. My conjecture about qualitative spatial reasoning is that it involves a vocabulary of PLACES whose relationships are described in symbolic terms. By PLACE, I mean a piece of space (point, line, region, volume, etc.) such that all parts of it share some property. The nature of a domain determines the notion of place appropriate to it. There might be more than one useful way to break up space even within a single domain, and the results of qualitative spatial reasoning must be integrated with other knowledge. This suggests embedding the place vocabulary in a more quantitative, analog representation.[1]

2.1 The Metric Diagram

We do not yet understand the complexities of human vision, so we do not know precisely what people compute from a diagram or how they do so. The role of a diagram can still be studied by building a representation that has some simple way of computing the answers to relevant questions. A type of geometry representation that I call a *Metric Diagram* was used in FROB to explore these issues. The geometric aspects of a problem are represented by symbolic elements whose parameters are numbers, embedded in a bounded global coordinate system. The vocabulary of elements required for the Bouncing Ball world consists of points, line segments, regions bounded by line segments, and pieces of vertically oriented parabolas. The mathematical simplicity of the elements and the availability of numerical parameters means analytic geometry can be used to calculate answers to most kinds of geometric questions.

In constructing the program three kinds of questions proved important. They will be called *identity, parity,* and *intersection.* Identity questions concern the relationship of the geometric elements with the descriptions of the objects whose geometric aspects they represent. Aside from being necessary for interpreting the results of the processes associated with the diagram, indexing the elements by what they represent can speed up searches that use the diagram. For example, detecting possible collisions with surfaces is much faster if only the surface geometries need to be tested against the trajectory as opposed to testing all elements in the diagram.

A geometric element divides space up into different pieces, which can be considered as "sides." Parity questions concern on what side of some element a

[1]By contrast, Hayes (see Hayes, 1979b) explicitly avoids the use of metric representations for space. I suspect that a metric representation will be required to make his concept of a history useful, in that to compare them requires having a common coordinate frame.

point is, and what sides of one element another is on. For example, to detect that a ball is placed inconsistently inside a solid requires being able to detect that the point which represents the ball at some point in time is inside the region that represents the solid.

Intersection questions are very important, because for physical things to interact they must "touch." They are answered by solving the equations attached to the elements to find possible points of contact and filtering the results with parity operations to account for the limited spatial extent of the elements. One use of intersection questions is finding out if a ball hits a particular surface, and if so, where.

2.2 The Space Graph

In FROB the *Space Graph* provides the vocabulary of places. Because all balls are point masses and are subject to the same forces, the Space Graph is independent of them and depends only on the surface geometry. Free space is divided into nonoverlapping regions in a way that simplifies the description of possible motions, as will be discussed in section 4. These regions and the edges that bound them are the Metric Diagram elements that form the nodes of the Space Graph. These nodes are connected by arcs that are labeled with the name of the relationship between them (such as LEFT or UP). Any other place required for

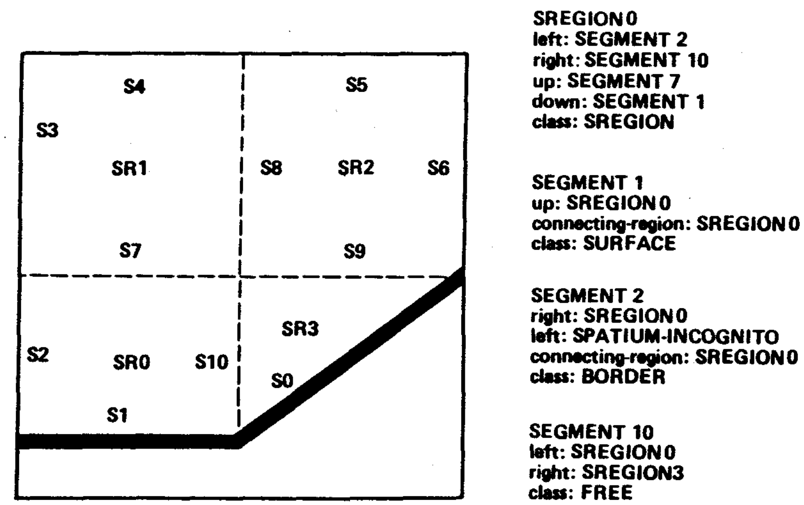


FIG. 4.2. Space Graph for a scene. The free space in the diagram is broken up into regions in a way that simplifies the description of the kinds of motion possible. The labels on the pointers indicate the spatial relationships between the nodes.

qualitative reasoning can be described by composing these places. The graph structure provides a framework for efficient processing (see sections 4 and 5). An example of the places in a scene and the graph structure they produce is contained in Fig. 4.2.

2.3 Comparison with Other Spatial Descriptions

The Metric Diagram has much in common with the descriptions used as targets for language translation of Waltz and Boggess (1979) and the imagery theory of Hinton (1979). It is quite different from the traditional "pure relational" geometric representations used in AI and the "naive analog" representations used by Funt (1976), and Kosslyn and Schwartz (1977). Both of these schemes are inadequate, but for different reasons.

Reasoning about space with just relational descriptions can be difficult. Transitive axioms such as

$$\text{Left-of}(X, Y) \land \text{Left-of}(Y, Z) \Rightarrow \text{Left-of}(X, Z)$$

are often needed to answer parity questions, and their use can lead to combinatorial searches, as pointed out in [Waltz and Boggess, 1979]. Relational systems are very weak models of space.[2] For a fixed vocabulary of predicates and relations there is only one full relational description (all possible relations and predicates are asserted) up to isomorphism between object names for any Metric Diagram, but for a relational description there can be infinitely many Metric Diagrams. In drawing a diagram from a relational description (as is often done in solving physics problems, for instance) the relational description must first be filled out and then actual parameters found which satisfy this description. The fact that people are willing to go through this trouble in generating pencil and paper diagrams seems to indicate that our fluency in dealing with space does not come solely from a set of very clever axioms for reasoning with a relational description.

The "Naive Analog" scheme uses an array to model space, representing the location and extent of an object implicitly by what cells contain symbols corresponding to that object. By explicit analogy with low-level vision, a simple local process called a "retina" is used to examine the array in order to compute answers to spatial questions. In Funt (1976) such a scheme was used to simulate falling blocks and in Kosslyn and Schwartz (1977) is the central feature of a theory of mental imagery. This representation has several flaws (aside from not corresponding to the facts available about retinal function). Putting the processing in the "retina" leads to performing searches to answer most questions. To place an object into an array requires choosing parameters for its location, scale and rotation and then turning on the correct cells in the array. The instantiation of a Metric Diagram element requires only the first part of this process, and since it can be used to answer the questions the array becomes superfluous. Hinton (1979) argues against the use of array based representations in mental imagery on the same grounds. For some geometric questions a fully parallel array scheme could have certain advantages, such as determining intersections in constant time. The tradeoffs involved in such a scheme, however, have yet to be determined.

[2]Hayes [Hayes, 1979a] notes that the axioms for the geometry of blocks in many problem solvers can be satisfied by modelling a block as an ordered pair of integers, one component for the number of blocks below it, and one component for discrete locations on the table. This is far from the intuitive notions of space they are intended to capture.

3. DESCRIBING A PARTICULAR MOTION

When people watch an object move, they generally couch their description in terms of a sequence of qualitatively distinct motion types. I call a network built from descriptions of motions linked by descriptions of the state of the object before and after each motion an *Action Sequence*.[3] The knowledge associated with each type of motion allows it to be further analyzed, the consistency of the proposed description to be checked, and permits making predictions about what will happen next. A drawn trajectory of motion in the Bouncing Ball domain and the schema of its associated Action Sequence is illustrated in Fig. 4.3.

The two basic types of motion in the Bouncing Ball world are flying and colliding. We denote occurrences of these motions by elements in the Action Sequence called FLY and COLLIDE. Flying up and flying down are separated into distinct acts because different things can happen after each of them. Acts that represent transitions to motion outside the domain of the Bouncing Ball world are CONTINUE for leaving the space enclosed by the diagram and AMBIGUITY-SLIDE/STOP, AMBIGUITY-SLIDE/STOP/FALL, and STOP when a ball interacts with a surface for any amount of time. Each act in the Action Sequence describes where it is occurring as well as when.

The description of a ball's state contains a parameter indicating the instant it applies to. It includes quantitative parameters such as the ball's position (specified by a point in the Metric Diagram), speed, and heading at that time. The kind of motion that will occur next, and what the ball is touching at that instant are other important parameters.

A description of the motion in terms of a qualitative state is also provided within the Action Sequence. The qualitative state of a ball includes the type of motion, position abstracted to a PLACE, and heading abstracted to a symbolic description, such as (LEFT UP). These qualitative states link the description of a

[3]The Action Sequence may be viewed as the *history* of a ball (in the Naive Physics sense of the term) since it contains explicit spatial and temporal limits.

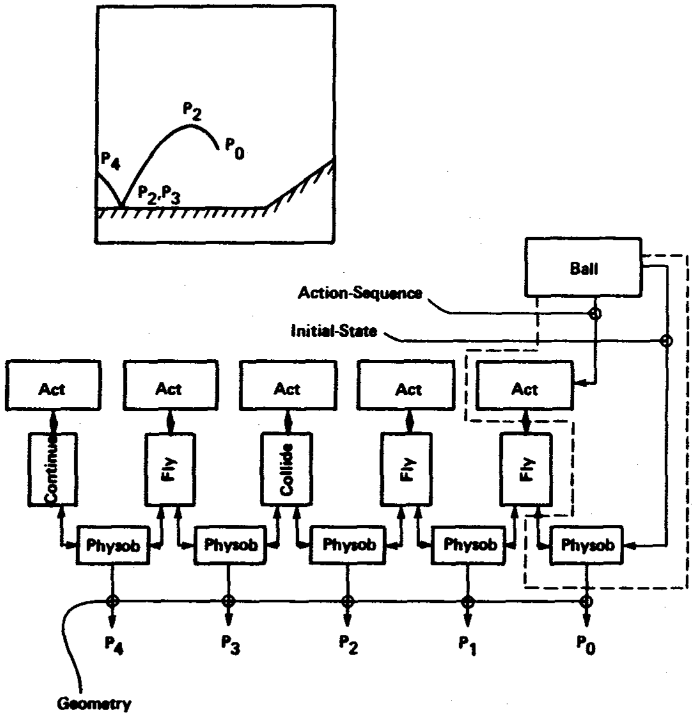


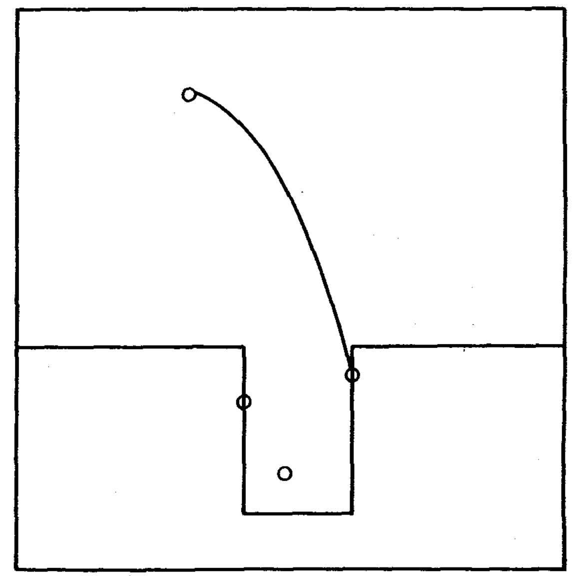
FIG. 4.3. Action Sequence Schema for Bouncing Balls. This schema describes the motion depicted in Fig. 4.1. The PHYSOB constraint describes the state of the ball at some instant in time, and the ACT constraints describe a piece of the ball's history.

particular motion to the description of possible motions explained in the next section.

In FROB the Action Sequence descriptions are embedded in a constraint language (see Steele & Sussman, [1978] for an overview, & Forbus [1981b] for a description of the particular language used). Each element of an Action Sequence is a constraint object, and they are connected together in a way such that partial information can be provided in whatever order is convenient. Local processes make deductions whenever possible, and can signal if an inconsistency is discovered.

The constraint descriptions used in FROB's Action Sequence include equations describing projectile motion to compute numerical values if numerical descriptions of the state parameters are obtained. The use of quantitative parameters in the qualitative description of motion makes possible a different kind of simulation from the usual incremental time simulations used in physics. When

numbers are provided, an Action Sequence can be produced by generating a description of the next motion from the last known state of motion. The time to generate the description, as well as the complexity of the result, depends on the qualitative complexity of the motion rather than some fixed increment of time used to evolve a set of state parameters. FROB's simulation capability was used as a way of generating motion descriptions for qualitative analysis.



```
 CONTRADICTION DISCOVERED CONCERNING (>>(A2 YSUM1 ENERGY F1)
 WHOSE VALUE 8.888888896 DEPENDS ON
 THE NEW VALUE 2.111111112 COMPUTED BY (RULE-3 . G0892) DEPENDS ON
1 ( >> Y   S1) = -4.0 from USER
 BOTH VALUES SHARE THESE ASSUMPTIONS-
2 ( >> (CHECKED-VALUE COR-CHECK S3)  (C-O-R  S3) ) = 0.5 from USER
3 ( >> Y- COMPONENT VELOCITY S5) = 0.0 from USER
4 ( >> Y S5) = 7.0 from USER
5 ( >> X S5) = -4.0 from USER
6 ( >> Y S3) = -3.0 from USER
7 ( >> X S1) = -2.0 from USER
8 ( >> X S3) = 2.0 from USER
  CHOOSE ONE TO RETRACT BY CALLING ANSWER WITH ITS NUMBER
 ;BKPT CONTRADICTION-HANDLER
```

FIG. 4.4. An inconsistent description of motion. This motion is impossible because the ball could not get as high as it does after the second collision unless it had gone higher on the first. If it had gone higher after the first, the second collision would not even have happened. To discover that this description is inconsistent FROB requires a specific velocity at the highest point and a specific value for the elasticity of the ball as well as the coordinates of the collision points.

A proposed motion can be analyzed by building an Action Sequence for it, letting the knowledge of the equations of motion attached to the resulting constraint network look for inconsistencies and fill in consequences of what is known. FROB's dependence on quantitative parameters in the Action Sequence is a drawback. For example, FROB can detect that the situation in Fig. 3.4 is inconsistent only after being given some final height for the ball and a value for the elasticity. People can argue that this proposed motion is impossible with simpler arguments that require less information. To deal with this, FROB could be extended with a more qualitative set of analysis methods. One such rule for the Bouncing Ball domain would be "A ball cannot increase its energy from one act to the next."

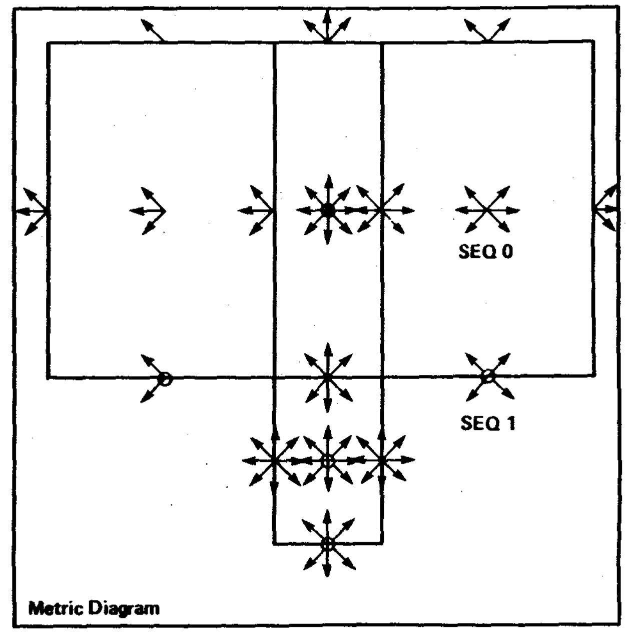
4. DESCRIBING POSSIBLE MOTIONS

There are some predictions people can make even when they know very little about a situation. For example, if a ball is bouncing leftwards on an infinite flat plane, it will never start going to the right unless something interferes with it. People can perform this and other inferences by the ability to describe the set of possible motions a ball might undergo.

One way to represent the possible motions of a ball is to use the idea of a qualitative state mentioned previously. Simulation rules can be written that operate on qualitative states, but because of the ambiguity in the description they may predict that several motions are possible from some state. There are only a small number of places and a small number of motions possible at each place, so all the possible kinds of motions can easily be computed. This process is called *envisioning*. The technique of envisioning was first introduced in deKleer, (1975) for answering simple questions about a scene directly and as a planning device for algebraic solutions to physics problems. In FROB the result of envisioning is called the *Sequence Graph*, which uses the Space Graph for its spatial framework (see Fig. 4.5). The Sequence Graph is more complicated because it deals with a truly two dimensional domain and includes the effects of dissipative forces. Like deKleer's envisioner it is used to answer simple questions directly, but is also used to assimilate global constraints on motion.

The place vocabulary of the Space Graph is chosen to keep the Sequence Graph simple yet precise. Each different surface and border of the diagram must be a place, but there are many possible ways to carve up free space. A non-overlapping decomposition is used so that any quantitative state will map to a unique qualitative state. The considerations that were imposed on the Space Graph by its role as a framework for the Sequence Graph are:

1. The motion description must be kept small. This means that unnecessary distinctions should be avoided.
2. The branching factor (the number of motions possible after some state)




```
— >>(what-is ( >> root sequence-graph phob))
(    ROOT SEQUENCE-GRAPH PHOB) = SEQ0
NIL
—>>(pseq seq0)
 THIS IS THE START NODE OF THE GRAPH FOR G2860
SEQ0
(FLY SREGION3 (LEFT DOWN))
 CAN BE REACHED BY (SEQ12)
 NEXT CAN BE (SEQ1 SEQ2)
SEQ0
—>>(pseq .seq1)
SEQ1
(COLLIDE SEGMENT9 (LEFT DOWN))
 CAN BE REACHED BY (SEQ0)
 NEXT CAN BE (SEQ3 SEQ4)
SEQ1
```

FIG. 4.5. A Sequence Graph. The arrows represent the direction of a qualitative state at the place the arrow is drawn. Circles represent states without well defined directions. The possible temporal orderings of the states are not depicted.

must be kept small. One way of achieving this is to have only a single place be reached when travelling in a particular direction from another place.

3. Thanks to gravity the simplest motion of the domain is bouncing up and down over a horizontal surface. To keep the description of this motion simple, cut space by vertical and horizontal lines.

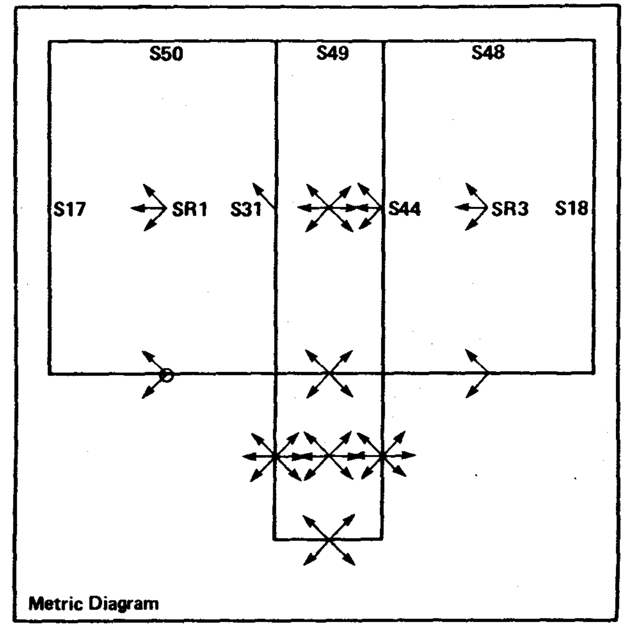
The Sequence Graph consists of all motions possible under the assumed initial condition. Knowing more about a ball than its state of motion at some time can restrict these possibilities. Energy limits the height a ball can reach, and knowing that a ball is perfectly elastic or completely inelastic excludes certain results of a collision. Assumptions about whether a ball must or may not reach a particular place or qualitative state can restrict the possibilities as well. The Sequence Graph can be modified by pruning states to reflect this information about the ball and its motion.

Each of the constraints above directly rules out some states of motion. The full consequences of eliminating such states are determined by methods that rely on specific properties of space and motion. It might appear that because this problem is concerned with belief revision, it could be solved by using a domain independent Truth Maintainence System (see Doyle, 1978; McAllester, 1980). A qualitative state would be "justified" if at least one predecessor state is possible and if one of the possible states after it is possible (unless it is a terminal state such as STOP). This does not work. The problem is that the Sequence Graph description contains a large number of cycles (corresponding to repetitive motion), making the computation of well founded support intolerably difficult. Instead the following facts of motion are used in pruning the Sequence Graph:

1. Only qualitative states that can be reached by some path of possible qualitative states from the initial one are themselves possible.
2. All motion occurs on a continuous path in space. Although implicit in (1) explicit use of this fact is advantageous because there are fewer places than qualitative states.
3. Unless a ball is perfectly elastic, it must either stop or leave the diagram.
4. A ball travelling in either the space above a surface with horizontal extent or between two surfaces with vertical extent in a particular direction can do so only if it either leaves the place going in that direction, changes direction, or stops within the place after moving in that direction.

Condition four is required to exclude situations whose qualitative description matches that of Zeno's paradox. An example is a ball bouncing on a horizontal surface travelling leftward, but never reaching the left border of the region it is in and never stopping.

Computations which use these facts are applied to the Sequence Graph to determine the consequences of the assumptions. Dependency information is




```
—>>(why-not seq37)
 (CONTINUE SEGMENT17 (LEFT)) IS UNATTAINABLE BECAUSE
 (CANNOT-REACH SEGMENT17)
SEQ37
—>>(why-not seq35)
 (CONTINUE SEGMENT50 (LEFT UP)) IS UNATTAINABLE BECAUSE
 (ENERGY)
SEQ35
—>>(why-not seq77)
 (CONTINUE SEGMENT18 (RIGHT)) IS UNATTAINABLE BECAUSE
 (REQUIRED-STATES (SEQ22))
SEQ77
—>>(pseq seq22)
SEQ22
 (PASS SEGMENT31 (LEFT UP))
  CAN BE REACHED BY (SEQ15)
  NEXT CAN BE (SEQ30
SEQ22
```

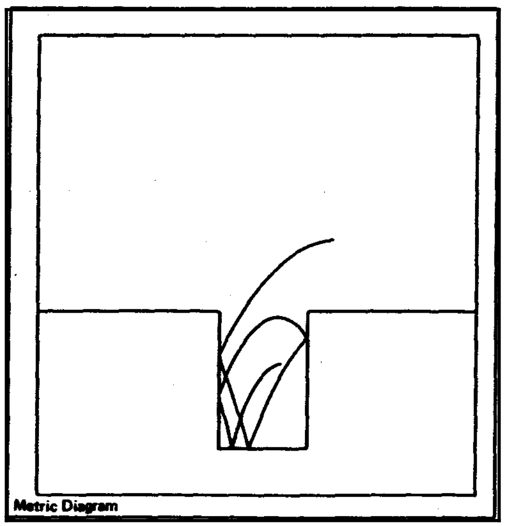
FIG. 4.6. Effects of assumptions on the Sequence Graph. Making assumptions about the physical properties of the ball or global properties of motion can reduce the ambiguity inherent in the Sequence Graph. Note the difference between this description and Fig. 4.5.

stored so that the effects of specific assumptions may be traced (see Fig. 4.6). Conflicting assumptions, overconstraint, and conflicts between a description of the actual motion (as specified by an Action Sequence) and its constrained possibilities are detected by FROB and the underlying assumptions are offered up for inspection and possible correction.

5. ANSWERING QUESTIONS

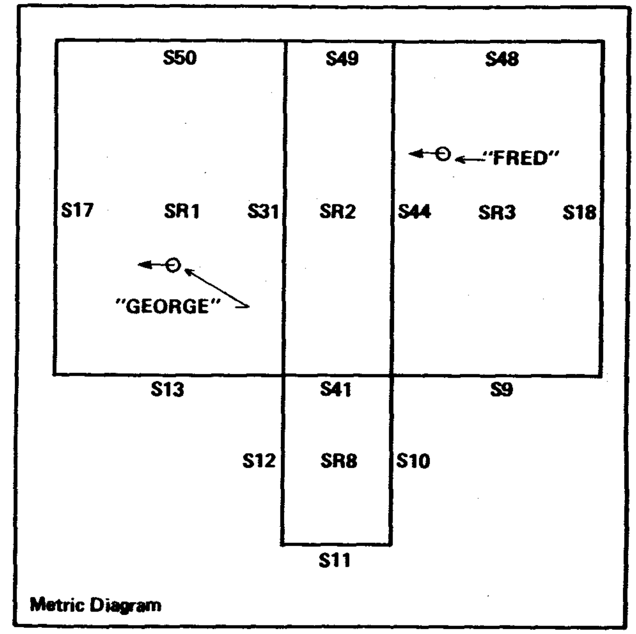
Many of the questions that could be asked by the Bouncing Ball domain can be answered by direct examination of the descriptions built by FROB. These include the first two questions in section 1.1. The three levels of motion description in FROB (the Action Sequence, the Sequence Graph, and the path of qualitative states corresponding to the Action Sequence) allow some kind of answer to be given even with partial information.

The more complicated questions about summarizing motion and collisions (questions 3 and 4 on p. 64) can be answered with additional computation. Summarizing motion includes determining whether or not a ball is trapped in a well (see Fig. 4.7), which can be done by examining a Sequence Graph for the last state in an Action Sequence to see if it is possible to be moving outside the places that comprise the well. The places where a ball can stop or leave the diagram are known, and so the possibilities for its final disposition can be described.




```
—>>(motion-summary-for b1)
 FOR G0364
 THE BALL WILL EVENTUALLY STOP
 IT IS TRAPPED INSIDE (WELL0)
 AND WILL STOP FLYING AT ONE OF (SEGMENT11)
NIL
```

FIG. 4.7. Summarizing motion.




```
—>>(collide? fred george)
(POSSIBLE AT SEGMENT50 SEGMENT17 SEGMENT13 SREGION)
—>>(cannot-be-at fred segment31)
(SEGMENT 31)
 UPDATING ASSUMPTIONS FOR ( >>INITIAL-STATE FRED)
 CHECKING PATH OF MOTION AGAINST ASSUMPT IONS
—>>(collide? fred george)
NO
—>>(what-is (>> state initial-state fred)
(   STATE INITIAL-STATE FRED) = (FLY (SREGION3) (LEFT))
NIL
—>>(what-is ( >> state initial-state george))
( >> STATE INITIAL-STATE GEORGE) = (FLY (SREGION) (LEFT))
NIL
```

FIG. 4.8. Collision problems.

Having several levels of detail allows collision questions to be answered more easily. Often a collision between two balls can be ruled out because the two balls are never in the same PLACE, as determined by examining their Sequence

Graphs. By relating the qualitative states for the Action Sequence with the time information associated with the ACTs, it is possible to determine whether or not a ball is in the same PLACE at the same time. With quantitative parameters in the Action Sequence description of motion it is possible to compute exactly where and when two balls collide if they do at all. Figure 4.8 contains the answers given by the program to collision questions in a simple situation.

6. DISCUSSION

6.1 Psychological Relevance

FROB captures the knowledge necessary to answer a number of the questions about the Bouncing Ball domain that people find easy to answer. This does not imply that it knows as much as people do about motion through space, nor that it uses what it knows in the same ways. Here we will examine some of these differences.

Two aspects of human understanding that are missing in FROB as a consequence of working only in a small domain can be thought of as *relevance* and *significance*. Relevance pertains to the uses of knowledge. A person uses knowledge about motion in free space to get around in the world—to avoid a falling rock, to throw a stone at something bothersome, etc. There is nothing inside FROB that corresponds to an explicit goal or value. FROB is an "it" only because in English it is convenient to characterize something that has processes and state as an entity.

By significance, I mean the ability to relate a piece of knowledge to other things you know. There is no interpretation of the tokens used in FROB's representations other than the processes that directly manipulate them, nor are they part of a larger corpus of knowledge. This makes FROB inflexible. For example, a person would understand that the only impact on his knowledge of halving the value of the gravitational constant would be that the value used in the equations of motion must be changed accordingly. He would understand that if gravity varied in magnitude with time the equations of motion would become more complex and if the sign varied as well his qualitative rules of motion would require revision.

Even within its intended domain, there are several differences between what FROB does and what people appear to do. First of all, people are far more flexible about the ways they divide space. In the well picture, for example, the chunks of space people usually describe are "inside the well" (SregionO and its borders) and "outside the well" (Sregional, Sregion2, Sregion3, and their borders). If necessary, they can make finer distinctions, such as which side of the well something is on (Sregionl or Sregion3). One way that FROB could be modified to exhibit this behavior would be to compute a tree of places, with the current Space Graph corresponding to the most detailed level of the tree. A set of rules about what level of the tree to use for envisioning and other computations

would be needed as well as methods to compare descriptions at different levels. Relaxing the restriction of a single qualitative description of space will almost certainly be necessary for more complex domains. This makes the existence of an underlying quantitative representation even more important, to serve as a communication device between different qualitative descriptions.

The answers people give when asked to describe what motions are possible in some situation do not look much like a natural language version of the Sequence Graph. Part of the lack of correspondence is due to the differences in place vocabulary described earlier, but not all. The descriptions people give often ignore certain of the possible motions. If pressed they can determine whether or not one of these unconsidered possibilities can indeed occur. There are several possible explanations for this phenomena. They might be performing envisionment, but pruning the results when communicating since it is very tedious to express a graph structure with many cycles as a string. They might instead be choosing only one alternative when using a set of qualitative simulation rules, and have the ability to backward chain using these rules when asked if some state that did not appear in their description can occur.

Another interesting question concerns how people actually assimilate qualitative assumptions about motion. In FROB a complicated process prunes the Sequence Graph and leaves behind a trace of why certain states cannot occur. A person may just perform the envisioning over again and stop when a state that is explicitly ruled out is reached. The states which in FROB would be pruned as a consequence of the new assumptions would then never be generated. In pure form this would not solve the problem of the Qualitative Zeno's Paradox, so some pruning would still be required. Careful protocols (and perhaps timing studies) of people deciding why or why not a state of motion is possible could shed light on the matter.

6.2 New Directions

Although the results of programs like FROB may be encouraging, there is still much to be understood about people's fluency in dealing with the physical world. Part of the progress can come from building programs like FROB, which reason about domains that form a separable part of physics (such as motion through space, sliding on a surface, etc.), by identifying a place vocabulary and a definition of qualitative state adequate for solving some set of problems. However, there are at least three areas which lie outside this approach and must be incorporated into it if we are eventually to succeed in creating a theory of common sense reasoning about physics.

The first area concerns the way qualitative knowledge is used, the *style* of reasoning performed. In FROB the use of qualitative knowledge centers around envisioning: the PLACE vocabulary was chosen to make it easy to perform, and creating and manipulating Sequence Graphs provides the means to answer questions that require qualitative knowledge. Much of the theory of the domain physics is encoded in the qualitative simulation rules and in the programs that prune the envisionment. Deductions based on the envisionment implicitly use the assumption that the simulation rules are complete and have been run to completion when ruling states in or out. Although envisioning is an important technique, I believe the burden of building a complete description of possible states is too onerous outside very small domains, and is too restrictive a style to capture all of the ways people use qualitative physical knowledge.

The envisionment for a complicated situation will be large because the PLACE description of the situation will be large. As discussed earlier, this could be ameliorated by greater flexibility in the PLACE vocabulary and especially by making it hierarchical. This leaves us with the issue of deciding what level of description to use in a problem (which is interesting on its own merits). A more serious complication arises if collisions between moving objects are explicitly represented. Because there is no information about time other than the orderings of the qualitative states a collision is possible after each state in the graph. This causes an explosive increase in the connectivity of the graph. A similar problem in more complex domains occurs when two processes or objects are acting in concert to produce an effect. An example is boiling water by passing steam through pipes in the container. To capture all possible states in this situation requires considering all possible time orderings for events, such as running out of water and shutting down the steam. Certainly some questions about these situations would require that much work, but surely not all. Still another complication ensues if analyzing a machine with controls. Consider for example a steam plant, which may have several hundred valves that could be adjusted at any time. There is no way to predict such an event with the physics of the plant, nor would explicitly representing the set of all the effects of all such possible events be attractive.

Even when envisioning is possible, it is not clear that people do it. Consider for example the situation in Fig. 4.9 which contains a lever with a pin in its path. When asked if the tip of the lever can get to point A (or B or C), the answer people give is no. The reasons they give are different in each case. For A the reason is that the pin stops the lever from moving. For B and C however, the reason is that the rigidity of the lever means it cannot bend, stretch, grow, or shrink to reach these points. The simplest envisioning system imaginable would only be able to say that these states are not considered possible. A more subtle one (as in FROB) would perform the envisioning without the pin, and prune the part of the path made impossible by the pin while keeping track of the reason for its rejection. But to remove the premise of rigidity would cause the number of possible motions to grow quite large. A simpler way to solve such a problem is to consider the ways the proposed state might come about. In this situation the only ways of getting the tip to B are by moving the pivot point or shrinking or bending the lever. These three prospects are easily eliminated by the assumptions that the pivot is fixed and the lever rigid, and so it can be concluded that the tip cannot be at B.

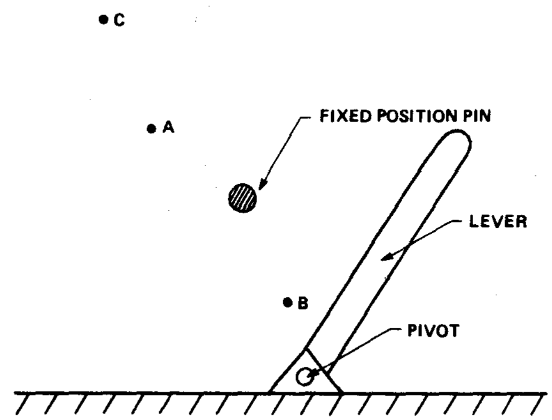


FIG. 4.9. Questions of excluded states. The pivot is in a fixed position and the lever is rigid. Consider whether or not the tip of the lever can reach A, B, and C in turn. The answers people give to these questions are the same that a system relying only on envisionment would give, but their reasons are quite different.

This example illustrates what can be done if qualitative knowledge is encoded more flexibly than in simulation rules. Questions about possible states can concern past states as well to infer possible causes, and in both cases simple deductions could often yield results. For example, a prerequisite for an object to get from one place to another is the existence of a path between the two places—no path, no motion. The same properties of physical theories that make envisioning possible, such as making influences explicit and few in number and allowing them to operate only through explicit connections,[4] make the number of theories that can link possible states small as well. This limited forward deduction is perhaps one of a number of "styles" of reasoning that should be explored.

Let us consider an abstract example to illustrate the other two areas I think are important for future work in this field. Imagine an object or a collection of objects for which we have a qualitative state description based on some theory about how they work. For the particular state they are in we know the possible states they might be in next, and perhaps if we have more detailed quantitative data we can actually calculate which of these possible states will occur. If we were to continue to project the possibilities for each possible state we would be performing envisioning. But these two possibilities do not exhaust our options. One thing we might do is make assumptions about the situation, based on past experience, that could lead to some of the alternatives being ruled out. We might

[4]In "Non-Naive" physics these properties are reflected in the unification of different influences into the notion of force and in the bias against action at a distance.

also consider what different things we would see depending on which state actually occurred, and just wait to find out what happens.

Both options remove us from the world of "armchair physics" in which our investigations so often reside. The main reason common sense physics is so interesting is that it is useful in helping us to get around in the world. If we are reasoning in order to deal with the world, we want to be able to make predictions quickly when possible. We want to reason about how some state of affairs could have come about so that it may be duplicated if desirable, or avoided if not. We need techniques to see if what we know about a situation is consistent with our physics for we can be mistaken, lied to, or ignorant. Current practice in artificial intelligence makes studying these kinds of issues difficult. A program is usually told all it will be told about a situation in one initial description and is not allowed to propose actions or execute actions that would provide more information. They are often designed to provide a quantitative answer in the style of a student solving a physics problem. These restrictions on the design of programs can be valuable simplifications of an already complex task, but if we maintain them too long our efforts may well become distorted.

Using experience probably has two roles in common-sense physics. It is sensible to assume that the description of objects we begin with when applying our theories about the world are fairly far from the idealizations of the physics. Experience with the world could guide the process of choosing the right physics and mapping from the given objects and relations to the idealizations. Another role for experience is the source of default assumptions. Few people viewing the lever (see Fig. 4.9) would volunteer that it would not reach B because the friction in the pivot was so high that the lever could not move at all, yet this could be the case. This aspect of physical reasoning skates close to the deep and turgid waters of learning and does not look simple.

Physical objects can be seen and touched, and while not all of the terms in our theories about them are perceptable (such as density), their effects certainly are. To discover if our theory about a situation is correct we are often goaded into performing actions upon the world. Few people, for example, would claim to fully understand an unfamiliar mechanical gadget just by looking at it; they push and poke its parts to see how it moves, and often modify their theories about it accordingly. A theory of common sense physical reasoning should include theories of what to observe and how to experiment within a situation. It should be able to deduce what the observable consequences of alternate theories concerning a situation are and deduce what sort of manipulations can be made to gain required information.

REFERENCES

Bundy, A. *MECHO: Year one*. Research Report No. 22, Department of Artificial Intelligence, Edinburgh, 1976.

de Kleer, J. *Qualitative and quantitative knowledge in classical mechanics*. Technical Report 352, MIT AI Lab, Cambridge, Mass., 1975.

de Kleer, J. *Causal and teleological reasoning in circuit recognition*. MIT AI-TR-529, Cambridge, Mass., Sept. 1979.

Doyle, J. *Truth maintenance systems for problem solving*. MIT AI Lab Technical Report 419, Cambridge, Mass., Sept. 1978.

Forbus, K. *A study of qualitative and geometric knowledge in reasoning about motion*. TR-615, MIT AI Lab, Cambridge, Mass., Feb. 1981. (a)

Forbus, K. *A Conlan Primer*. BBN Technical Report, 1981. (b)

Funt, B. V. *WHISPER:A computer implementation using analogues in reasoning*. Unpublished doctoral dissertation, University of British Columbia, 1976.

Hayes, P. J. The naive physics manifesto. In D. Michie (Ed.), *Expert Systems in the Micro-Electronic Age*, Edinburgh University press, May 1979. (a)

Hayes, P. J. *Naive physics I—ontology for liquids*. Memo, Centre pour les etudes Semantiques et Cognitives, Geneva, 1979. (b)

Hinton, G. Some demonstrations of the effects of structural descriptions in mental imagery. *Cognitive Science*, Vol. 3, No. 3, July–September 1979.

Kosslyn, S. & Schwartz, S. P. A simulation of visual imagery. *Cognitive Science*, Vol. 1, No. 3, July 1977.

Marr, D. *Artificial intelligence—a personal view*. Memo No. 355, MIT AI Lab, Cambridge, Mass., March 1976.

McAllester, D. *An outlook on truth maintenance*. MIT AI Lab Memo 551, Cambridge, Mass., August 1980.

McDermott, J., & Larkin, J. *Re-representing textbook physics problems*. Presented at the 2nd National Conference of the Canadian Society for Computational Studies of Intelligence, Toronto 1978.

Novak, G. *Computer understanding of physics problems stated in natural language*. Technical Report NL-30, Computer Science Department, The University of Texas at Austin, 1976.

Steele, G., & Sussman, G. *Constraints*. Memo No. 502, MIT AI Lab, Cambridge, Mass., November 1978.

Waltz, D., & Boggess, L. Visual analog representations for natural language understanding. Presented at *IJCAI-79*, Tokyo, Japan, August 1979.