SPATIAL AND QUALITATIVE ASPECTS OF REASONING ABOUT MOTION

Kenneth D. Forbus
MIT Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Mass. 02139

ABSTRACT

Reasoning about motion is an important part of common sense knowledge. The spatial and qualitative aspects of reasoning about motion through free space are studied through the construction of a program to perform such reasoning. An analog geometry representation serves as a diagram, and descriptions of both the actual motion of a ball and envisioning are used in answering simple questions.

I Introduction

People reason fluently about motion through space. For example, we know that if two balls are thrown into a well they might collide, but if one ball is always outside and the other always inside they cannot. The knowledge involved in this qualitative kind of reasoning seems to be simpler than formal mechanics and appears to be based on our experience in the physical world. Since this knowledge is an important part of our common sense, capturing it will help us to understand how people think and enable us to make machines smarter. The issues involved in reasoning about motion through space were studied by constructing a program, called FROB, that reasons about motion in a simple domain. I believe three important ideas have been illustrated by this work:

1. A quantitative geometric representation simplifies reasoning about space. It can provide a simple method for answering a class of geometric questions. The descriptions of space required for qualitative reasoning can be defined using the quantitative representation, making it a communication device between several representations of a situation

2. Describing the actual motion of an object can be thought of as creating a network of descriptions of qualitatively distinct types of motion, 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 description of the kinds of motion possible from some state (called the envisionment) is useful for answering certain questions and for checking an actual motion against assumptions made about it. The assimilation of assumptions about global properties of the motion into this description makes heavy use of spatial descriptions.

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. We ignore the exact shape of balls, motion after two balls collide, spin, motion in a third spatial dimension, air resistance, sliding motion, and all forces other than gravity.

The initial description of a situation is a diagram containing a description of the surfaces and one or more balls, as in figure 1.

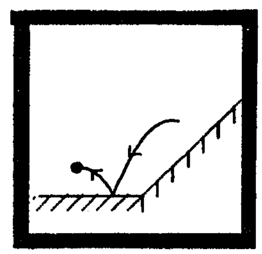

Fig. 1. A typical scene from the Bouncing Ball world
A situation in the Bouncing Ball World consists of a diagram that specifies surfaces and one or more balls. This drawing only shows the geometric aspects of the descriptions involved.

When given a description of a situation in the Bouncing Ball world, 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 its descriptions to accommodate this information, complaining if inconsistencies are detected. Questions may be asked by calling procedures that interrogate these descriptions. The four basic questions FROB can answer are: (1) What can it (a ball) do next?, (2) Where can it go next?, (3) Where can it end up?, and (4) Can these two balls collide?

II 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 the former may be found in [1], while the sheer complexity of algebraic manipulations argues against the latter. I conjecture that the fluency people exhibit in dealing with space comes mainly from using their visual apparatus. One example is the use of diagrams. 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.

We do not yet understand the complexities of human vision, but the techniques of analytic geometry can be used to provide decision procedures for geometrically simple cases. FROB uses a Metric Diagram, which is a representation of geometry that combines symbolic and numerical information. The geometric aspects of a problem are represented by symbolic elements whose parameters are numbers in a bounded global coordinate system. The representation is used to answer questions about basic spatial relationships between elements, such as which side of a line a particular point lies or whether or not two lines touch. Calculation based on properties of the elements suffices to answer these questions.

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. In FROB the Space Graph provides the vocabulary of places. Since 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 in the Metric Diagram. Free space is divided into regions in a way that insures the description of qualitative state (described below) will be unique and simple, and these regions and the edges that bound them are the nodes of the Space Graph. These nodes are connected by arcs that are labelled with the name of the relationship between them (such as LEFT or UP). Any other place required for qualitative reasoning can be described by composing these nodes, and the graph structure provides a framework for efficient processing (see section 4). An example of the places in a scene and the graph structure they produce is contained in figure 2.

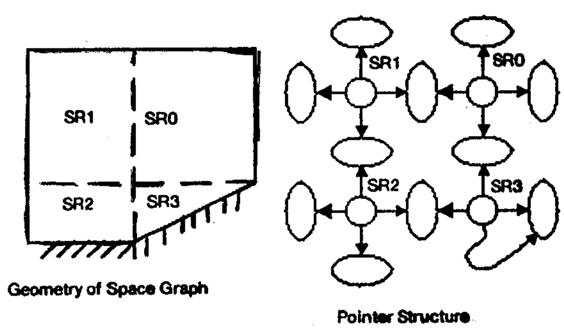

Fig. 2. Space Graph for a scene

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 which indicate the spatial relationships between the nodes are not shown due to lack of space.

III Describing a Particular Motion

When we watch an object move, we generally couch our description in terms of a sequence of qualitatively distinct motion types. We will 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. 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 figure 3.

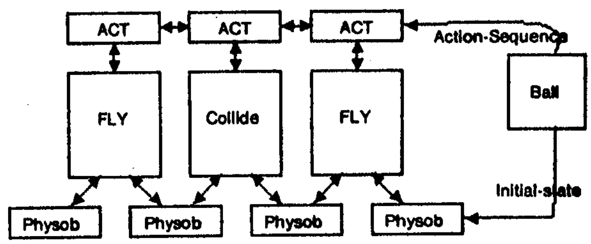

Fig. 3. Action Sequence Schema for Bouncing Balls

This schema describes the motion depicted in Figure 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.

The two basic types of motion in the Bouncing Ball world are FLY and COLLIDE. The difference in computing boundary conditions between flying up and flying down requires their consideration as separate acts in the sequence, and additional motion types are defined for transformations to motions outside the Bouncing Ball world (such as CONTINUE for leaving the diagram and SLIDE/STOP when a ball is travelling along a surface). The description of a ball's state includes such information as its velocity (quantitative if known, or just in terms of a rough heading like (LEFT UP)) and what it is touching.

In FROB the Action Sequence descriptions are embedded in a constraint language (see [2] for an overview), and 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.

Simulation is not the only way an Action Sequence can be created. A network of constraints can be built to describe some proposed motion and the knowledge of the equations of motion can be used to analyze it to see if it is consistent. The dependence on quantitative parameters in FROB's analysis is a drawback. For example, FROB can detect that the situation in figure 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.

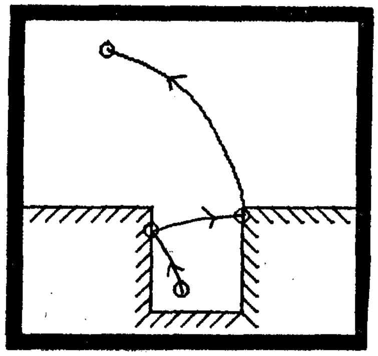

Fig. 4. An inconsistent description of motion

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.

The basic idea of an Action Sequence seems highly suited as a target representation for parsing quantitative data about motion, perhaps gleaned by perception. For this purpose a more qualitative set of methods for analysis would have to be encoded. An example of such a rule for the Bouncing Ball domain would be "A ball cannot increase its energy from one act to the next".

IV Describing Possible Motions

The quantitative state of a ball consists of its position and velocity. A notion of qualitative state can be defined which generalizes position to be a PLACE, the velocity to be a symbolic heading (such as (RIGHT DOWN)), and makes explicit the type of motion that occurs. A set of simulation rules can be written to operate on qualitative state descriptions, but because of the ambiguity in the description the rules may yield several motions possible from any given state. Since there are only a small number of places and a small number of possible qualitative states at each place, all the possible kinds of motion from some given initial qualitative state can easily be computed. This description is called the envisionment (after [3]) for that state. deKleer used this description to answer simple questions about motion directly and plan algebraic solutions to physics problems.

In FROB envisioning results in the Sequence Graph, which uses the Space Graph for its spatial framework (see Figure 5). It is used for summarizing properties of the long term motion of an object, evaluating collision possibilties, and assimilating assumptions about the global properties of motion. Only the assimilation of assumptions will be discussed here.

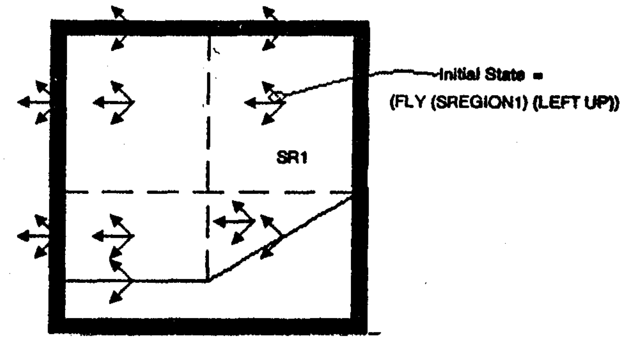

Fig. 5. A Sequence Graph

The arrows represent the direction of a qualitative state at the place the arrow is drawn. Circles represent states without well defined directions. The pointers expressing the possible temporal orderings of the states are not shown.

Knowing more about a ball than its state of motion at some time can restrict the kinds of motion possible to it. 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 possibilties 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. Among these properties are the fact that a motion of an object must be "continuous" in its state path (which means that the active part of a Sequence Graph must be a single connected component) and that the space it moves in must be connected (which is useful because there are many fewer places than qualitative states in any problem). Dependency information is stored so that the effects of specific assumptions may be traced. Conflicting assumptions, overconstraint, and conflicts between a description of the actual motion (as specified by an Action Sequence) and its constrained possibilties are detected by FROB and the underlying assumptions are offered up for inspection and possible correction.

V Answering Questions

Many of the questions that could be asked of the Bouncing Ball domain can be answered by direct examination of the descriptions built by FROB. These include questions (1) and (2) above. The three levels of description of motion 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.

More complicated questions (such as (3) and (4) above) can be answered with additional computation using these descriptions. Determining whether or not a ball is trapped in a well (see figure 6) can be done by examining a Sequence Graph for the last state in an

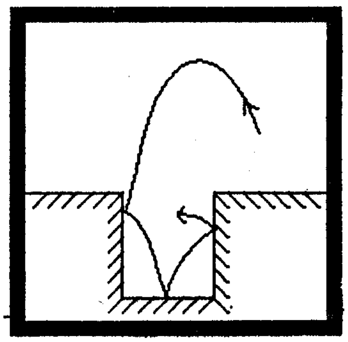

Fig. 6. Summarizing Motion

->>(motion-summary-for b1)
FOR B1
THE BALL WILL EVENTUALLY STOP
IT IS TRAPPED INSIDE (WELL0)
AND IT WILL STOP FLYING AT ONE OF (SEGMENT11)

Action Sequence to see if it is possible to be moving outside the places that comprise the well. 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. With the Action Sequence description of motion it is possible to compute exactly where and when two balls collide if they do at all. Figure 7 contains the answers given by the program to collision questions in a simple situation.

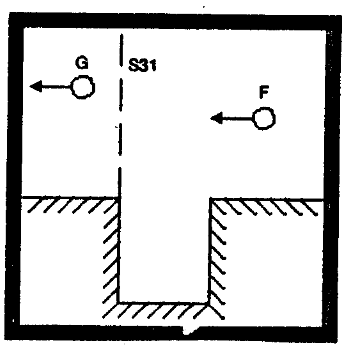

Fig. 7. Collision Problems

->>(collide? f g)
(POSSIBLE AT SEGMENT13 SREGION1 ...)
->>(cannot-be-at f segment31)
(SEGMENT31)
UPDATING ASSUMPTIONS FOR (>> INITIAL-STATE F)
CHECKING PATH OF MOTION AGAINST ASSUMPTIONS
->>(collide? f g)
NO

VI Relation to Other Work

The focus of this work is very different from that of [4][5][6], which are mainly concerned with modeling students solving textbook physics problems. All of the problems dealt with in these programs were static, and the representation of geometry expressed connectivity rather than free space. Issues such as getting algebraic solutions and doing only the minimal amount of work required to answer a particular question about a situation were ignored here in order to better deal with the questions of spatial reasoning and the semantics of motion.

The process of formalizing common sense knowledge is much in the spirit of the Naive Physics effort of Hayes (described in [7]). The Action Sequence, for example, may be viewed as the history for a ball since it contains explicit spatial and temporal limits. However, this work is concerned with computational issues as well as issues of representation. Unlike this work, Hayes (see [8]) explicitly avoids the use of metric representations for space. I suspect that a metric representation will be required to make the concept of a history useful, in that to compare them requires having a common coordinate frame.

The Metric Diagram has much in common with the descriptions used as targets for language translation of [1] and the imagery theory of [9]. Arguments against the traditional "pure relational" geometric representations used in AI and the "naive analog" representations used by [10],[11] may be found in [12].

The concept of envisioning was first introduced in [3] as a technique for answering simple questions about a scene directly and as a planning device for algebraic solutions. The inclusion of dissipative forces, a true two dimensional domain, interactions of more than one moving object, and its use in assimilation of global constraints on motion are the envisioning advances incorporated in this work.

VII Bibliography

[1] Waltz, D. and Boggess, L. "Visual Analog Representations for Natural Language Understanding" in Proc. IJCAI-79, Tokyo, Japan, August 1979

[2] Steele, G. and Sussman, G. "Constraints" Memo No. 502, MIT AI Lab, Cambridge, Massachusetts, November 1978

[3] deKleer, Johan "Qualitative and Quantitative Knowledge in Classical Mechanics" Technical Report 352, MIT AI Lab, Cambridge, Massachusetts, 1975

[4] Bundy, A. et. al. "MECHO:Year One" Research Report No. 22, Department of Artificial Intelligence, Edinburgh, 1976

[5] 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

[6] McDermott, J. and Larkin, J. "Re-representing Textbook Physics Problems" in Proc. of the 2nd National Conference of the Canadian Society for Computational Studies of Intelligence, Toronto 1978

[7] Hayes, Patrick J. "The Naive Physics Manifesto" unpublished, May 1978

[8] Hayes, Patrick J. "Naive Physics 1:Ontology for Liquids" unpublished, August 1978

[9] Hinton, G. "Some Demonstrations of the Effects of Structural Descriptions in Mental Imagery" Cognitive Science, Vol. 3, No. 3, July-September 1979

[10] Funt, B. V. "WHISPER:A Computer Implementation Using Analogues in Reasoning" PH.D. Thesis, University of British Columbia, 1976

[11] Kosslyn & Schwartz, "A Simulation of Visual Imagery" Cognitive Science, Vol. 1, No. 3, July 1977

[12] Forbus, K. "A Study of Qualitative and Geometric Knowledge in Reasoning about Motion" MIT AI Lab Technical report, in preparation.