

# Imagery for Open-World Spatial Problems

Boi Faltings

Swiss Federal Institute of Technology (EPFL)  
1015 Lausanne, Switzerland

Pearl Pu

University of Connecticut  
U155, Storrs, CT 06269  
USA

May 15, 2002

## 1 Modeling Spatial Properties

In her paper ([4]), Janice Glasgow addresses many important advantages that imagery models can offer over sentential models such as predicate logic. In this response, we point out another important advantage of imagery with respect to the *formulation of models*.

Consider the problem of reasoning about placing two-dimensional polygonal “blocks” in a space, illustrated in Figure 1. For purposes such as planning, knowledge about the blocks world is commonly formalized as rules which define the legal operations. Rules are based on symbolic features, for example, a rule may state that:

$$HOLDING(x, t) \wedge CLEAR(y, t) \wedge PUTON(x, y, t) \Rightarrow ON(x, y, t + 1)$$

meaning that executing the operation *PUTON* on  $x$  and  $y$  at time  $t$  will result in  $x$  being on top of  $y$  at the following instant, under the condition that at time  $t$ , the manipulator was holding  $x$  and  $y$  was clear.

However, as illustrated by Figure 1, such a rule may create an incorrect result when other objects are present: placing  $B$  on top of  $A$  results in an

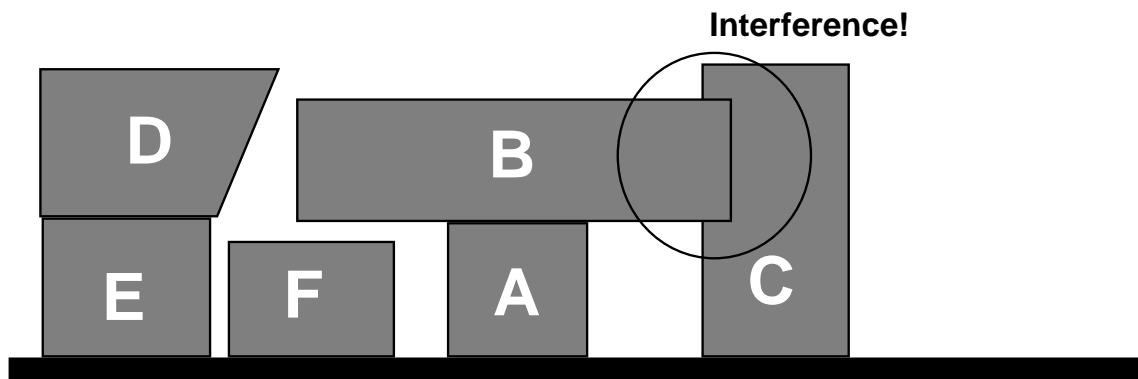


Figure 1: *Emergent features in the blocks world: placing B on A may be possible by a local rule, but interferes with blocks C and D which are in the vicinity of A. The interference between B and C is an emergent feature which arises only in this situation.*

interference with block  $C$ . To correct the rule, we should introduce a feature such as *INTERFERES* which explicitly avoids such interference:

$$HOLDING(x,t) \wedge CLEAR(y,t) \wedge PUTON(x,y,t) \wedge \neg(\exists z)INTERFERES(z,x,y,t) \Rightarrow ON(x,y,t+1)$$

But how could a feature such as *INTERFERES* be defined? It obviously depends on the shapes and positions of  $x$ ,  $y$  and  $z$ . But more precisely, it must take into account the *combination* of *all* objects that might interfere with placing  $x$  on  $y$ : block  $C$  would not interfere with  $B$  if block  $D$  were not in the position it is in. Interference cannot be expressed as a local property of a small number of objects, but *emerges* through the interaction of potentially *all* objects in the scene.

Modeling spatial interaction in predicate logic requires formulating an amount of knowledge which grows exponentially in the number of objects in the situation. This makes logic impractical for any problem of realistic size. However, it is indeed possible to compute the interference property on the basis of an imagery model: the space available for placing  $y$  is defined as the intersection of two regions:

- the region where  $y$  would be on top of  $x$ , and
- the complement of the regions taken up by the other objects.

This operation is illustrated in Figure 2, where the two regions are shown in grey. Block  $B$  can be placed on  $A$  whenever the intersection of the regions is large enough to include  $B$ ; in this case, it is just large enough by a small margin. “Imagery” in this context does not have to mean a direct analogical representation, but could also refer to transformed spaces. For example, transforming physical space into *configuration spaces* is practical in many applications such as robot path planning ([5]) or mechanism kinematics ([2]). A configuration space transformation would also simplify the solution of Figure 2: instead of testing whether  $B$  can fit into a region, we only have to test whether a region is non-empty.

Note that while the computation of regions and their intersections can be formulated in logic, this would amount to using logic to simulate an imagery operation and not as a sentential knowledge representation. Imagery is thus indispensable for modeling and solving spatial reasoning problems.

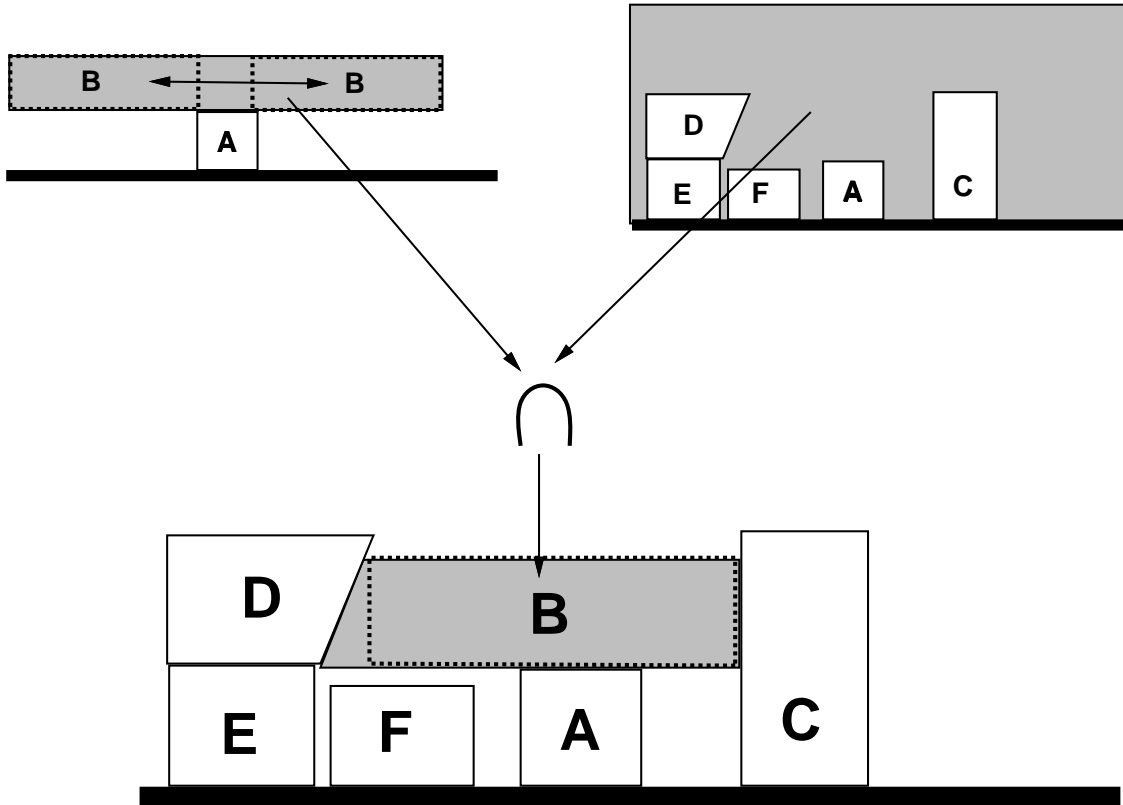


Figure 2: *Emergent properties can be readily obtained by imagery operations. Here, the question of whether B can be placed on top of A is resolved by testing whether the intersection of two regions is big enough to contain B.*

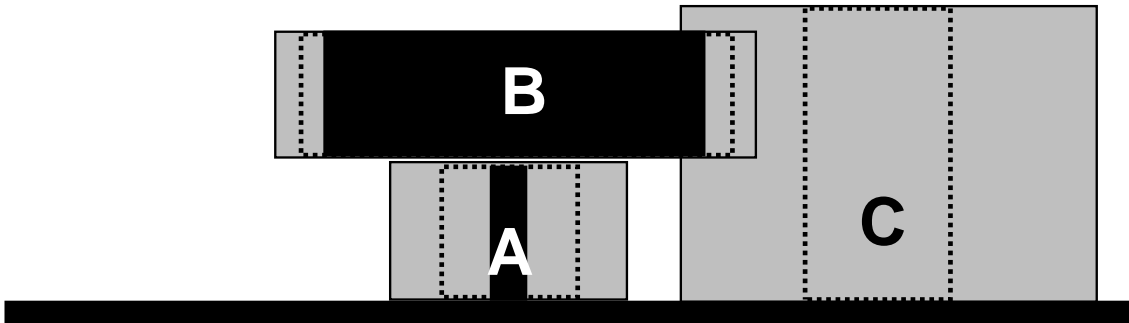


Figure 3: *Representing uncertainty about the positions of blocks by core regions (black) and fuzz regions (grey).*

## 2 Representing Vagueness: A Challenge for Imagery

An imagery representation of spatial problems such as the blocks world has an important drawback: a situation must be *precisely* represented. There is no provision for the partial or qualitative descriptions such as they arise in planning. In robot motion planning, these have usually been incorporated by adding dimensions to the space which reflect additional degrees of freedom - but this is not a cognitively plausible solution.

Solutions to this problem are sometimes possible if we are willing to allow an image to contain generalized objects which do not necessarily correspond to physically existing ones. As an example, consider again the blocks world. If we are uncertain about the position of a block, we could represent it by two regions:

- A *core* region which is definitely completely occupied by the block, no matter what its precise position is.
- A *fuzz* region of positions which might be occupied by the block, but only when it is in certain positions.

A query whether a certain position is legal can now have three answers:

- true when there is no overlap between fuzz or core regions.

- unknown when there are overlaps, but no overlaps between two core regions.
- false when there is an overlap between two core regions.

An example of a representation with such characteristics is the fuzzbox of Davis ([1]).

Depending on the problem, one can distinguish *vagueness*, where the description leaves a freedom of choice, and *uncertainty*, where this choice will be fixed by nature. In the case of vagueness, the representation can be made more precise by adding assumptions: for example, in Figure 3 one could assume that the blocks are in the positions shown by the dashed lines. In the case of uncertainty, the minimal condition for guaranteeing a correct solution is to replace the objects by their fuzz regions. In this way, additional degrees of freedom can be represented in images without increasing their dimensionality. Exploring such techniques is a promising direction for increasing the computational power of imagery representations.

### 3 A practical example: Spatial Planning

Problems of spatial reasoning occur in many problems of practical interest. In one of our research projects, we have constructed a program which solves a spatial planning problem by means-ends analysis ([3]). An example of a problem solved by the program is shown in Figure 4.

People have no difficulty with such tasks, but techniques developed in research on robotic motion planning run into severe complexity problems when solving it. This complexity is due to the fact that they miss the focus which means-ends analysis provides to people. Using means-ends analysis in automatic motion planning, however, requires a qualitative symbolic world model. Due to the problem described in Section 1 of this paper, formulating this world model in sentential representations leads to combinatorial explosion and thus again to unmanageable complexity.

The approach we have adopted is one where the features of the qualitative world model are generated incrementally as they are required by the planner. The state of the world is represented analogically using imagery. Planning proceeds by backward chaining and at each stage requires operators which achieve relevant goals by moving blocks. We represent goals, motions



and obstacles as regions in an imagery model. Operators are generated by manipulating these regions. For more details on the process, the reader is referred to [3].

One feature of our planner which is particularly interesting is the way in which we qualitatively represent motions. A qualitative path is a maximal set of topologically equivalent paths passing between the same obstacles. In planning, we need to *protect* a path from being clobbered with obstacles. It turns out that the protections required to ensure the existence of a topological path can be represented as particular regions such that whenever an obstacle intersects both regions, the path is blocked, and otherwise it is feasible. This makes it possible to represent the constraints imposed by a *set* of qualitatively equivalent motions by *one* single precise region. We think that similar expressions of vagueness or uncertainty could be useful in other problems as well.

## References

- [1] **Ernest Davis**: “Representing and Acquiring Geographic Knowledge,” Pitman Publishing, London, 1986
- [2] **Boi Faltings**: “Qualitative Kinematics in Mechanisms,” *Artificial Intelligence* **44**(1), 1990
- [3] **Boi Faltings, Pearl Pu**: “Applying Means-Ends Analysis to Spatial Planning,” Working notes of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations, Stanford, 1992.
- [4] **Janice I. Glasgow**: “The Imagery Debate Revisited: A Computational Perspective,” *Computational Intelligence*, this issue.
- [5] **T. Lozano-Perez, M.A. Wesley**: “An Algorithm for Planning Collision-free Paths among Polyhedral Obstacles,” *Comm. ACM* **22**(10), 1979