Chapter X

A Commonsense Approach to Representing Spatial Knowledge Between Extended Objects

Tiansi Dong
Cognitive Ergonomic Systems, Germany

ABSTRACT

This chapter proposes a commonsense understanding of distance and orientation knowledge between extended objects, and presents a formal representation of spatial knowledge. The connection relation is taken as primitive. A new axiom is introduced to govern the connection relation. Notions of ‘near extension’ regions and the ‘nearer’ predicate are coined. Distance relations between extended objects are understood as degrees of the near extension from one object to the other. Orientation relations are understood as distance comparison from one object to the sides of the other object. Therefore, distance and orientation relations are internally related through the connection relation. The ‘fiat projection’ mechanism is proposed to model the mental formation of the deictic orientation reference framework. This chapter shows diagrammatically the integration of topological relations, distance relations, and orientation relations in the RCC frameworks.

INTRODUCTION

When we open our eyes, we see a snapshot view of the spatial environment. In this chapter, spatial environments are those in which objects are projectively as large as, or larger than, the body but can be visually apprehended from a single place without appreciable locomotion (Montello, 1993, p. 315). They are vista spatial environments following Montello (1993), or the space surrounding the body following (Tversky, Morrison, Franklin, & Bryant, 1999; Tversky, 2005).

We subjectively decompose snapshot spatial environments into objects and spatial relations among them; recognize them, describe spatial relations, and identify whether it is the target environment which we want to enter, even detect object movements in the environment. For example, when you have the first snapshot view of your office in the morning, you will recognize objects in the room, such as a chair and a desk which are indistinguishable from yours, describe their relations, such as the chair is near and in front of the desk, identify whether
it is your office, and detect object movements, such as the chair has been moved a bit to the left than its location when you left yesterday. There are several interesting issues involved in the knowledge that people have about snapshot views of spatial environments.

Firstly, from snapshot views, we only see part of objects, and other parts may be blocked. However, we can recognize them. This can be evidenced by the Gestalt Theory, e.g., (Koehler, 1929; Koffka, 1935; Wertheimer, 1958), and research in object recognition, e.g., (Biederman, 1987; Spelke, 1990; Buelthoff & Edelman, 1992; Humphreys & Khan, 1992; Tarr, 1995; Tarr & Buelthoff, 1995). We can even recognize objects in snapshot views both from the real world and from the virtual world, e.g., films, TV programs, and photos. When you open your office door in the morning, you receive the light reflection from the chair, and you recognize your office chair; when you see a photo of your office, you receive the light reflection from the photo, and you also recognize your office chair. Recognizing objects either in the real environment or in the photo owes to the light reflection and to our recognition activity—we have knowledge about objects. The knowledge of objects resides in the memory and is activated either by some external stimuli or by certain mental desires. Secondly, object recognition means categorization. Objects in the same category are considered equivalent. Rosch, Mervis, Gray, Johnson, & Boyes-Braem (1976) argued that categories of recognized objects are structured such that there is generally one level of abstraction at which we find it easiest to name objects and recognize them the fastest, namely the “basic level category”. The basic level is the first categorization made during perception of the environment. This suggests a link between our spatial knowledge and our spatial descriptions. That is, the spatial knowledge acquired through perception structures our language, i.e., Tversky & Lee (1999).

Thirdly, from the perspective of cognitive informatics (CI), i.e. Wang (2003), the meta-cognitive process layer carries out the fundamental and elementary cognitive processes commonly used by processes in higher cognitive layers. It is a critical layer which connects with the perception layer and all other higher layers. How the brain builds spatial representations in the meta-cognitive process layer has become a topic of great interest in cognitive informatics, i.e., Wang, Wang, Patel, and Patel, 2006. The aim of the chapter is to present internal relations between three kinds of spatial relations. The reminder of the chapter is structured as follow: we first propose a commonsense understanding of the spatial knowledge that can be acquired through perception; then we review the Region Connection Calculus (the RCC-8 theory) in the literature and point out one weakness in its axioms; after that we set the starting point of the formal representation, propose a new axiom to govern a characteristic property of the connection relation and present the formalism of distance knowledge between regions; then we show the formalism of the orientation knowledge between regions, and present two examples to demonstrate the integration of distance and orientation relations into the RCC framework.

A COMMONSENSE UNDERSTANDING

Suppose you are standing at the entrance door shown in Figure 1 (a) and looking into the room, you could recognize not only objects in the room, but also spatial relations among them. You would recognize spatial relations between yourself and objects, such as you are nearer to the balloon than to the writing-desk, and also spatial relations among objects, e.g., the balloon is in front of the writing-desk. If you turn on a flashlight, it will emit a light beam. Imagine our eyes are such a flashlight that they can emit a light beam and that we see the side of the object which blocks the light beam. Then, that an object is in front of the observer can be understood as follows: If the observer faces to the object, there will be a light beam which connects with both the eyes and the object, shown in Figure 1 (b); that an object is at the left-hand side of the observer can be understood as: If the observer turns to the left, there will be a light beam which connects with both the eyes and the object, or if the observer faces forward, there will be a light beam which connects with the left side of the observer’s face and the object—The observer can prove this by turning to the left to see whether there is a light beam connecting the object and the eyes; that the observer is nearer to object A than to object B can be explained as: The light beam which connects with the observer and object A is shorter than the light beam which connects with the observer and object B.

Imagine the observer does not emit light beams, rather ultrasonic waves, like blind bats, or imagine a blind man with a stick, they can also know whether there are obstacles in front of them and which obstacle is nearer to them than the other, if they know the distance that the emitted ultrasonic wave travels or the extent reached