

Operational Representation – A Unifying Representation for Activity Learning and Problem Solving

Seng-Beng Ho

Temasek Laboratories, National University of Singapore
5A Engineering Drive 1, #09-02, Singapore 117411
hosengbeng@nus.edu.sg

Abstract

A typical AI system engages many levels of cognitive processing from learning to problem solving. The issue we would like to address in this paper is: Can a unified representational scheme be used in learning processes as well as the various levels of cognitive processing from concept representation to problem solving including the generation of action plans? In a previous paper we defined a set of representations called “atomic operational representations” that employs an explicit representation of the temporal dimension and that can be used to ground concepts in the physical world, such as concepts that involve various activities and interactions. In this paper we apply operational representations in a unified manner to the following cognitive processes: 1) the unsupervised learning and encoding of causal rules of actions and their consequences; and 2) the application of the learned causal rules to problem solving processes that produce desired action plans. The unique and explicit temporal characteristic of operational representations is the key feature that allows the encoded concepts to be used in a unified manner across the various levels of cognitive processing. Hence, abstractions in the form of operational representations have an important role to play in AI.

Introduction

A typical AI system engages many levels of cognitive processing from learning to problem solving. The issue we would like to address in this paper is: Can a unified representational scheme be used in learning processes as well as the various levels of cognitive processing from concept representation to problem solving including the generation of action plan?

In a previous paper (Ho 2012) we defined a set of temporally explicit representations called “atomic operational representations” that can be used to ground concepts in the physical world, such as concepts that involve various activ-

ties and interactions. These representations are usable both for the recognition and generation of activities and interactions. The key idea behind operational representations is the explicit representation of the temporal dimension. In this paper we apply operational representations in a unified manner to various cognitive processes. Three main issues are addressed: 1) the representation of actions and action consequences in the form of operational representations; 2) an unsupervised causal learning process that allows a cognitive system to learn the rules that represent the causes and effects of actions; and 3) the application of the learned causal rules to problem solving processes that produce desired action plans.

The unique and explicit temporal characteristic of operational representations is the key feature that allows the encoded concepts to be used in a unified manner across the various levels of cognitive processing. Hence, abstractions in the form of operational representations have an important role to play in AI. The representational framework was developed as part of the effort to address a grand challenge to computational intelligence as posed in Ho 2013 and Ho and Liausvia 2013a.

Review of Related Work

A lot of research has been carried out in the area of computer vision in which the information in the visual world is processed for the purposes of object recognition, scene description and activity classification (e.g., Ali and Shah 2010; Shapiro and Stockman 2001; Szeliski 2010, Yuan et al. 2011) but these efforts have not emphasized characterizing and representing the conceptual aspects of the visual world based on the activities and interactions that are observed to take place between various entities in the world, and how causal rules encapsulating the concepts of these activities and interactions can be learned and extracted from visual observations and then used subsequently for

problem solving. These are the main issues we are addressing in the current paper. In the area of qualitative physics (Hobbs and Moore 1985; Liu and Daneshmend 2004; Weld and de Kleer 1990; etc.), it has been shown how causal rules can be encoded for the purpose of physical reasoning but these rules are built-in and not learned from visual observations.

Psychologists and linguists have understood for some time that perception is important for conceptualization (e.g., the work by Miller and Johnson-Laird 1976) but no effort has been forthcoming in these fields in terms of detailed computational characterizations of the ideas involved. In recent years, cognitive linguistics has arisen as a new paradigm for the study of meanings through perceptual and cognitive characterizations (e.g., the work of Evans and Green 2006; Geeraerts 2006; Langacker 2008, 2009; etc.). This brings us closer to establishing a link between perception and conception but detailed computational characterizations are still currently lacking in cognitive linguistics. In cognitive linguistics, it has been shown how an explicit representation of time is essential to capture the meaning of events and processes (e.g., Langacker 2008). This is in agreement with our operational representational scheme but in this paper we develop the idea further in a computational way by showing how this “correct” representation of the “ground meaning” of various concepts using an explicit temporal representation can participate in learning and problem solving.

Representations of Interactions with Explicit Causal Agent

A causal agent of activities and interactions can be something that is visible and/or tangible such as a point of light (that can cause, say, an amoeba to move toward it), an object (that can, say, push another object), a person’s hand, or something that is intangible such as an electric or magnetic “force.” In this section we describe the representations of some fundamental, atomic, and elemental causes of activities and interactions within the operational representational framework (Ho 2012) using one dimensional (1D) spatial situations.

Propulsion and Materialization

In our daily experiences, we identify “forces” as causal agents of movements whatever their origin and physical nature. In this paper, we represent a force and its associated direction in the form of a thick arrow that is present in the same spatiotemporal location as the elemental object that the force is influencing as shown in Figure 1(a). Figure 1(a) is basically a 2-dimentional *spatiotemporal* map of a 1-dimensional *spatial* movement in the spirit of the explicit temporal representation/operational representation as artic-

ulated in Ho (2012). (Note that the elemental “blobs” - the filled circles - that represent “elemental” objects in Figure 1 is an “elemental” bit of occurrence relevant to the cognitive task at hand - the “blob” could correspond to an atom in the physical world, a vehicle or a person, or the recognizable points on a human body or on the leaves of a tree.) We assume these forces can be observed and identified directly by a visual perceptual system if it is visible or indirectly through reasoning by observing the consequential changes in the state of the world (there will be more about cause identification in a later section). In all cases in this paper we would consider only forces of one unit magnitude that will cause an elemental object to move one elemental spatial location in one elemental time frame but our representational scheme is general enough to represent forces of all kinds of magnitude.

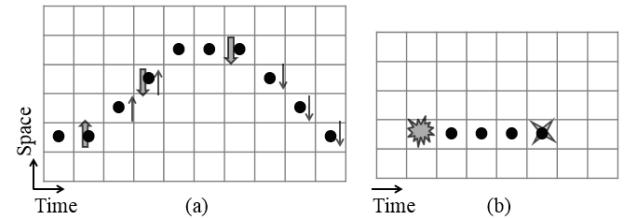


Figure 1. Representation of causal agents and their effects. (a) A force acting on an object. (b) A Materialization action (“starburst”) resulting in a persistent object (filled circle) and the subsequent Dematerialization action (“four-pronged star”).

In Figure 1(a) we show the object acquiring a momentum (represented by a thin arrow that accompanies the object) after being moved by the force and continuing to move until its movement is counteracted by an opposing force (after which the momentum disappears and the object comes to a halt). This kind of dynamics takes place when either the object involved is traveling in free space with no friction or the force is large relative to the weight of an object or relative to the friction on the ground. Otherwise, a force will have to be continually applied to the object in every time frame to sustain its movement and the object will halt when the force ceases. The observation that a momentum exists in an object (and hence the momentum labeling in the same square as the object) is derived from the fact that it continues to move with no apparent cause(s)/force(s) acting on it.

In Figure 1(b) we show the operations of Materialization and Dematerialization. An operation of Materialization is to bring something into a location seemingly from nowhere and we represent the action with a “starburst” icon as shown in Figure 1(b). The action of Materialization is applied at a spatiotemporal location with no object in it and an object would appear in the next time instance at the

same location. A point of light that is being switched on and appears at an elemental spatiotemporal location is also a kind of materialization.

An opposite action to that of *Materialization* is *Dematerialization* – the disappearance of an object (or switching off of something like a point of light) from a spatiotemporal location. This is shown in Figure 1(b) as a “four-pronged star.” Right after the action of *Dematerialization*, the object ceases to exist in the next elemental time step.

In our previous paper (Ho 2012) we mentioned that for a visual cognitive system to observe and represent activities and interactions such as those described above, it must have an experiential memory to buffer the changes in the temporal dimension so that the elemental/atomic changes that characterize these activities and interactions can be captured in temporally explicit operational representations.

Reflection, Obstruction, and Penetration

As objects move around in the environment, very often they would come into contact with each other or affect each other through forces acting from a distance and interactions would take place, changing the nature of their original movement. In Figure 2 we show a situation in which a mobile elastic object (a filled circle) reflects off an immobile elastic object (a filled square). On the left of Figure 2 we depict the reflection in 1D space and the spatiotemporal representation of the event is shown on the right. The impinging mobile object is shown to have a momentum toward the stationary object initially and after the contact event, its momentum reverses direction.

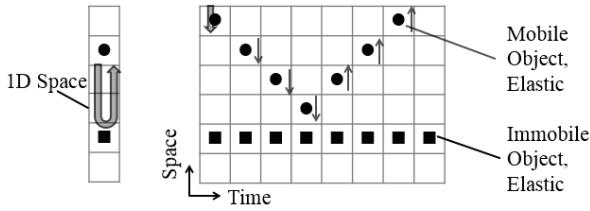


Figure 2. Representation of Reflection.

In Figure 2 a conceptual hierarchy can be built up to represent the concept of *Reflect* based on the atomic operators *Move-Down*, *Contact* and *Move-Up* as described in Ho 2012.

Figure 3(a) illustrates a situation in which both the impinging object and the immobile object are inelastic (represented by an unfilled circle and an unfilled square respectively) and the immobile object absorbs the momentum of the impinging object completely. Hence, after the contact, the impinging inelastic object comes to rest next to the immobile inelastic object. Figure 3(a) shows that the impinging object originally carries a momentum with it and

after the contact event with the immobile inelastic object the momentum disappears.

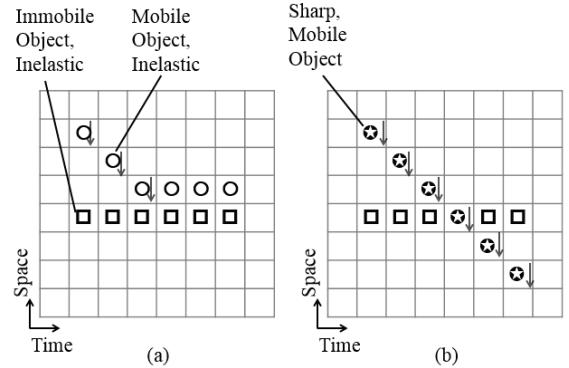


Figure 3. (a) Hit-and-stay event, no reflection. (b) A Penetration event.

In Figure 3(b), we illustrate the situation in which a “sharp” incoming object (represented by a circle with a star inside) is able to *Penetrate* an immobile object because of its “sharpness” property. The spatiotemporal picture of penetration is that the incoming object will not reflect nor will it come to rest next to the immobile object. It will continue its motion unimpeded to a point beyond the immobile object.

Attach, Detach, Push, and Pull

So far our discussions involve elemental objects that move independently in various types of interactions. Figure 4(a) shows the operation of *Attach* in which two elemental objects are joined together so that subsequently they would move in unison. We use an “attach-link” between the objects to indicate that they are attached after the *Attach* action which is represented as two opposing arrows. Figure 4(b) is the reverse operation *Detach* that would remove the attach-link and subsequently the elemental objects may not always move in unison. In the real physical world, an explicit attach-link may not be visible for objects that are attached together but a “Pull Test” as described below would determine if two objects are indeed attached together. The link is a representation of the state of the objects in the same vein as the momentum arrow in, say, Figure 1(a).

Figure 4(c) shows that when a force is applied to a first object in a situation in which a second object is in contact with (and not necessarily attached to) it and the force is applied in the direction of the second object, the first object then ends up “pushes” the second object and causes it to move. This is a *Push* operation. The *Push* operation also applies to the situation in which both the objects involved are attached. Also, in Figure 4(c) we show that both objects

(the pushing and the pushed objects) acquire a momentum after the *Push* action.

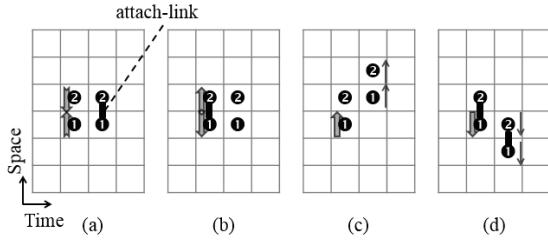


Figure 4. (a) Attach and (b) Detach actions. (c) Push event. (d) Pull event. The “attach-link” indicates attachment between the elemental objects.

Figure 4(d) shows a *Pull* action in which the force is applied on one of the two objects in the direction away from the second object. In this case, both the first and second objects must be attached to allow the *Pull* action to take effect and when it does, both objects would move in unison as shown. In the real world, if two objects are next to each other and perhaps touching each other but there are no explicitly visible linkage between them, the only way to establish whether the two objects are attached is to execute a *Pull* action on one of them. If the second object moves in unison with the first one, then the *Attach* relationship exists between the two objects. This is a “*Pull Test*” for attached objects.

Unsupervised Causal Rule Extraction and Problem Solving

Causal rules specify the consequences of taking actions, be it on the part of the cognitive system or some external agents. Knowing the causal rules of physical behaviors, which are like the “mental models” of the world, the cognitive system can use them for reasoning. The system can begin with a current state of the world and reason with these rules what might happen in the future in a forward reasoning process. The system can also use these rules for problem solving – how could a sequence of actions be assembled to achieve a certain given goal. In this section we describe how causal rules can be learned from the environment and how they can be used for problem solving.

Unsupervised Causal Rule Extraction

In a complex environment, there could be multiple causes for any given event and a sophisticated process of causal analysis is needed to disentangle the causes involved before an accurate mental model of the world can be constructed (Pearl 2009). In the various elemental activities and interactions that we describe in Figures 2 to 4, there

are only single causes involved and hence the unsupervised learning process is relatively simple. We relegate the method for disentangling and uncovering causes in complex situations to a future paper. Our focus here is to illustrate how operational representations can be used for causal rule representations and for subsequent problem solving processes.

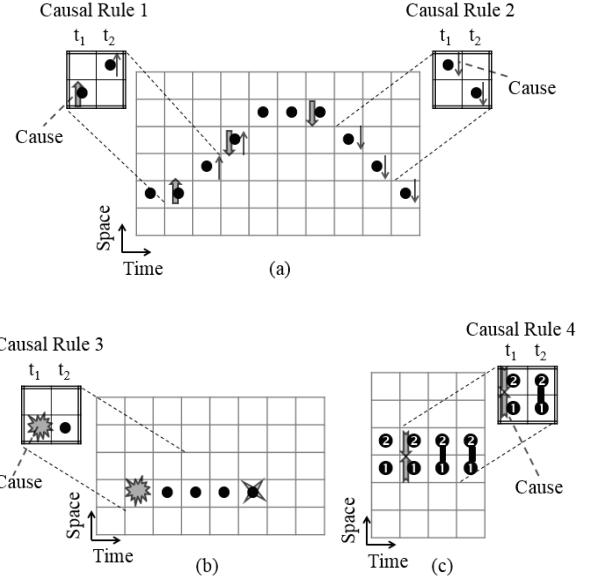


Figure 5. First stage of causal rule encoding. (a) Extraction of causal rules for “force” and “momentum.” (b) Causal rule for “Materialization action.” (c) Causal rule for “Attach action.”

In Figures 5 and 6 we depict a two-stage method of unsupervised causal learning in which the entire spatiotemporal causal structure is firstly “cut-out” from the activity in the environment (Figure 5). This is the “cookie-cutter” unsupervised learning approach used in Uhr and Vossler (1981) for *spatial* pattern recognition generalized to process *spatiotemporal causal* patterns here. And then in the second stage (Figure 6), the *cause* of the activity in the extracted causal rule is identified.

Figures 5(a), 5(b) and 5(c) contain the same activities as in Figures 1(a), 1(b) and 4(a) respectively and these activities have been explained above in connection with these figures. The extracted Causal Rule 1 in Figure 5(a) encodes the knowledge that “if at time frame 1 (t_1) a force is applied to the object in an upward direction, then the object will move to the next elemental location in the same direction at time frame 2 (t_2) and acquire a momentum in that direction.” The rule is extracted not in any verbal or propositional form but in the same spatiotemporal form as the activity or interaction itself, and in the next section – the Problem Solving section – we show how this direct encoding can facilitate the problem solving process. Causal

Rule 2 in the same figure encodes the knowledge that “if at t1 a momentum accompanies an object in a downward direction then at t2 it will move an elemental step in that direction and continue to be accompanied by a momentum in the same direction.” These rules are then stored in a Causal Rule Base and used in subsequent reasoning and problem solving processes that will be described in the next section.

Figures 5(b) and 5(c) likewise show the extraction of causal rules that encode the knowledge of the “*Materialization action*” (Causal Rule 3) and the “*Attach action*” (Causal Rule 4) respectively. Many other rules can likewise be extracted from the activities and interactions in other parts of the events in Figures 1(a) and 1(b), in the other events in Figures 3 (*Hit-and-stay* and *Penetration*) and 4 (actions of *Detach*, *Push*, and *Pull*), and in other events or situations that the cognitive system may come across.

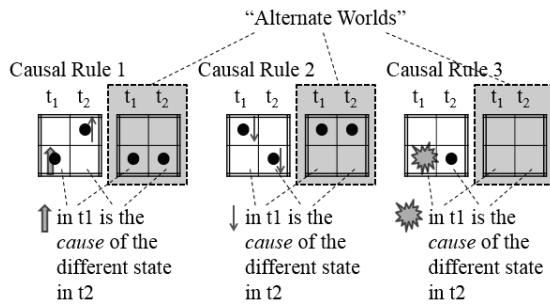


Figure 6. Second stage of causal rule encoding: Causes identified based on alternate world/counterfactual comparisons.

In the second stage of the causal learning process, the causes in the 4 causal rules in Figure 5 (namely the “force,” the “momentum,” the “*Materialization action*,” and the “*Attach action*” respectively of Causal Rules 1, 2, 3, and 4) are identified using a simplified “alternate world” or counterfactual analysis as shown in Figure 6. Basically, something is the cause of certain changes if the world or objects in the world behave differently “had it not been there.” – so, say, for Causal Rule 1, had the “force” not been there the object would have stayed in the same position in the next time frame in the alternate world shown. Again, in principle, to uncover causes in the general situation is not easy even if there are only single causes and it usually requires a long period of observation but here we take the “cookie cutter” approach and assume that if something appears to be a cause when first observed, assume that it is unless otherwise informed by future evidence. Consider the case of Causal Rule 1 in Figure 6. There is an “alternate world,” which was an observation made at a different spatiotemporal location (and presumably stored in some episodic memory of the

cognitive system so that it can be retrieved for comparison) in which the upward force was not present and the object involved did not move (comparing time frame t1 in both worlds and time frame t2 in both worlds). This is taken as a counterfactual indication that the force is the cause of the movement and this is indicated in Causal Rule 1 accordingly. Similarly for the other rules.

Problem Solving

For problem solving, the cognitive system begins with a start state and a goal state, and reason out, through a search process, either in the forward direction (starting from the start state to the goal state) or the backward direction (from the goal state to the start state) the sequence of actions that can be used to go from the start to the goal state. In the Causal Rule Base described above that was created by the process of unsupervised learning of causal rules, each causal rule captures an elemental action that can be chained together by the problem solving process into a sequence of actions to provide a solution to the problem.

We illustrate this with an example. In Figure 7, the “Problem” states that starting from a physical configuration which consists of two mobile objects (*Objects 1* and *2*) and an intervening *immobile* object, a sequence of actions is to be found such that sometime later (the indefinite interval represented by the long vertical gray bar) *Object 1* and *Object 2* acquire some upward momentum, and at that time instance, *Object 1* can be in any position specified by the gray bar under it. (Note that the problem statement is in a form of a pattern to be matched to the rules in the Causal Rule Base.)

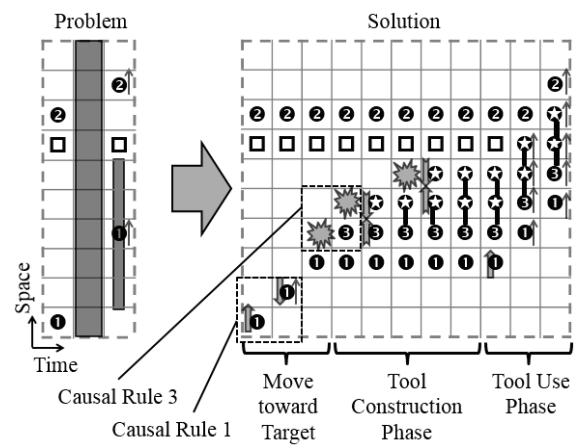


Figure 7. A problem in which the solution – the action plan - is to construct a tool to penetrate a certain immobile object (see text for explanation).

Starting from the start state, the reasoning process considers the various actions/causes (by scanning through all

the rules in the Causal Rule Base) that may be performed on *Object 1*, say, and create a search space to find a solution to the goal. The search space could potentially be very large and a method called Incremental Causal Rule Chunking could be used to reduce it, the details of which is relegated to a separate paper (Ho and Liausvia 2013a; Ho and Liausvia 2013b). Due to the lack of space, we will not show the search process in detail but just the results of the search. A possible “Solution” is shown on the right side of Figure 7. The solution shows that *Object 1* first moves toward *Object 2* as a result of an upward force applied to it and then it halts 2 spatial locations later as a result of a counteracting force. Then, a tool is constructed by materializing one mobile object and two mobile “sharp” objects (the processes of which were discussed earlier in connection with Figures 1(b) and 3(b)) over a few time frames and attaching them together in such a way that the sharp end of the tool can then be used to penetrate the immobile object to push *Object 2*.

In Figure 7, to avoid clutter, we show only the involvement of Causal Rules 1 and 3 in certain parts of the action plan construction process. We can identify the Tool Construction Phase and the Tool Use Phase of the solution shown in Figure 7 in many real world problem solving processes as well – e.g., building a car (constructing a “tool”) and driving it to somewhere (using the tool). In the real world, “materialization” of various objects necessary for the construction of a tool could correspond to bringing parts from other places to the scene of the construction of the tool. Those processes would require other subproblems to be solved.

In Figure 7, the attachment of the 3 subparts of the tool is not necessary for the pushing action of *Object 1* to propel them toward *Object 2*, but would be necessary if later *Object 1* were to “retrieve” the tool by pulling it “backward/downward.” A program implementation of this unsupervised learning and problem solving process is described in Ho and Liausvia 2013a and 2013b.

Discussion and Conclusions

In addition to what we illustrated in our previous paper (Ho 2012) that operational representations with an explicit temporal dimension seem a natural and intuitive way to represent many concepts of activities and interactions, in this paper we further showed that operational representations can provide a unifying framework across various levels of cognitive processing from unsupervised causal learning of activities and interactions to problem solving. It should not be a surprise that operational representations have such a unifying property. Activities and interactions in the physical world by their very nature connote a temporal dimension, and action rules and action plans also by

their very nature connote a temporal dimension. Thus, when the temporal dimension is captured explicitly in a representation, the representation has the cognitive unifying properties. Hence, abstractions in the form of operational representations have an important role to play in AI.

References

- Ali, S., and Shah, M. 2010. Human Action Recognition in Videos using Kinematic Features and Multiple Instance Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (2): 288-303.
- Evans, V., and Green, M. 2006. *Cognitive Linguistics: An Introduction*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Geeraerts, D. ed. 2006. *Cognitive Linguistics*. Berlin: Mouton de Gruyter.
- Ho, S.-B. 2012. The Atoms of Cognition: A Theory of Ground Epistemics. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*, 1685-1690. Austin, TX: Cognitive Science Society.
- Ho, S.-B. 2013. A Grand Challenge for Computational Intelligence – A Micro-Environment Benchmark for Adaptive Autonomous Intelligent Agents. *Proceedings of the IEEE Symposium Series on Computational Intelligence: Intelligent Agents*, 44-53. Piscataway, New Jersey: IEEE Press.
- Ho, S.-B., and Liausvia, F. 2013a. Knowledge Representation, Learning, and Problem Solving for General Intelligence. *Proceedings of the Sixth International Conference on Artificial General Intelligence*, 60-69. Berlin: Springer-Verlag.
- Ho, S.-B., and Liausvia, F. 2013b. Incremental Rule Chunking and Problem Solving. *Proceedings of the 1st BRICS Countries Conference on Computational Intelligence*, in press. Brazil.
- Hobbs, J. R., and Moore, R. C. eds. 1985. *Formal Theories of the Commonsense World*. Norwood, New Jersey: Alex Publishing.
- Langacker, R. W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Langacker, R. W. 2009. *Investigations in Cognitive Grammar*. Berlin: Mouton de Gruyter.
- Liu, J., and Daneshmend, L. K. 2004. *Spatial Reasoning and Planning: Geometry, Mechanism, and Motion*. Berlin: Springer-Verlag.
- Miller, G. A., and Johnson-Laird, P. N. 1976. *Language and Perception*. Cambridge, Massachusetts: Harvard University Press.
- Pearl, J. 2009. *Causality*. Cambridge, UK: Cambridge University Press.
- Shapiro, L. G., and Stockman, G. C. 2001. *Computer Vision*. New Jersey: Prentice Hall.
- Szeliski, R. 2010. *Computer Vision: Algorithms and Applications*. Berlin: Springer.
- Uhr, L., and Vossler, C. 1981. A Pattern-Recognition Program that Generates, Evaluates, and Adjusts Its Own Operators. In Feigenbaum, E. A., and Feldman, J. eds. *Computers and Thought*. Malabar, Florida: Robert E. Krieger Publishing Company, Inc.
- Weld, D. S., and de Kleer, J. eds. 1990. *Readings in Qualitative Reasoning About Physical Systems*. San Mateo, California: Morgan Kaufmann Publishers, Inc.

Yuan, J., Liu, Z., and Wu, Y. 2011. Discriminative Video Pattern Search for Efficient Action Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33 (9): 1728-1743.