Issues in Temporal and Causal Inference

Pei Wang[1] and Patrick Hammer[2]

[1] Department of Computer and Information Sciences, Temple University
1925 North 12th Street, Philadelphia, PA 19122, USA
[2] Institute of Optimization and Discrete Mathematics, Graz University of Technology
Steyrergasse 30, 8010 Graz, Austria

Abstract. This paper discusses several key issues in temporal and causal inference in the context of AGI. The main conclusions are: (1) the representation of temporal information should take multiple forms; (2) classical conditioning can be carried out as temporal inference; (3) causal inference can be realized without a predefined causal relation.

A central function of intelligence is *prediction*, the ability for a system to anticipate future situations according to past experience. It is often considered as a form of *temporal inference* or *causal inference*. This paper focuses on several key issues in this type of inference, by introducing the approach taken in NARS (Non-Axiomatic Reasoning System), and comparing it with other approaches.

NARS is an AGI system designed according to the theory that *intelligence* is the ability for a system *to adapt to the environment while working with insufficient knowledge and resources*. The system takes the form of a general-purpose reasoning system, and carries out various cognitive functions (learning, planning, decision making, etc.) in a unified process. The theory and its formal model are described in [31, 32], as well as in other publications. Limited by the length of the paper, in the following only a small part of the system is described.

1 Integrated Representation of Temporal Information

NARS uses a formal language *Narsese* to represent various types of knowledge:

Term. A term names a concept in the system. In its simplest form, an atomic term is just a unique identifier, such as *bird* for the concept "bird".

Compound Term. A compound term is composed from other terms by a connector. For example, $([yellow] \cap bird)$ is a compound term for "yellow bird".

Statement. A statement is a compound term representing the substitutability of one term by another one. For example, "Tweety is a yellow bird" is represented by statement "$\{Tweety\} \rightarrow ([yellow] \cap bird)$". A statement with a truth-value measuring its evidential support is called a *judgment*.

Event. An event is a statement whose truth-value is specified for a duration. For example, "Tweety is following Bob" is represented in Narsese as "$(\{Tweety\} \times \{Bob\}) \rightarrow follow$", and the statement is *true* in the period when Tweety is following Bob, but neither before nor after that period.

Operation. An operation is an event that can be realized by the system itself. For example, "to follow Bob" is represented in Narsese as operation "$\Uparrow follow(\{Bob\})$", which the system can realize by directly executing it.

The formal definitions of the symbols used above are given in [32], and here they only need to be intuitively understood. Also, for the current discussion, it is enough to see the memory of NARS as a collection of interrelated concepts.

In this way, NARS uniformly represents all empirical knowledge as sentences in a formal language, while still keeps the differences among types of knowledge. This design is very different from the tradition of cognitive architectures, where the common practice is to distinguish "semantic/declarative memory", "episodic memory", and "procedural memory" from each other, and to handle them in separate modules, each with its storage structure and processing mechanism [17, 14, 6]. There have been other attempts to unify these memory modules, such as in a *graphical* model [25], while NARS does it in a *logical* model that has some similarity with logical programming [13], even though the memory of NARS can also be roughly seen as a conceptual graph.

Since an event is just a statement whose truth-value is specified for a period, the most straightforward representation of temporal information is to attach a time interval to each event [1, 18], or even to every statement, since accurately speaking, every conceptual relation hold in an interval, including "forever" as a special case. NARS does not take this approach, because in different situations the accuracy in specifying the beginning and ending of an event varies greatly, so to use a single unit of time by which all events are measured is probably neither necessary nor possible for an AGI. To be natural and flexible, in NARS an event can be seen as both a point and an interval in time, depending on the desired granularity. This treatment is consistent with the opinion that "The unit of composition of our perception of time is a *duration*" [15]. Therefore, the temporal information of an event is specified *relatively* with respect to another event, using one of the two built-in temporal relations: *sequential* and *parallel* (also known as *before-after* and *at-the-same-time*), which correspond to the *precedes* and *overlap* predicates in the Russell-Kamp construction [15].

As a reasoning system, NARS runs by repeating a working cycle, and in each cycle the system carries out a step of inference, as well as some simple input/output activities. Just like a biological system uses certain rhythmic event as a "biological clock", NARS uses its working cycles as an internal clock, since each working cycle roughly takes a short constant amount of time. Using this internal clock, NARS can express the durations of certain events. For example, it can represent something like "Event A is observed, then, after 5 cycles, event B is observed", where the "5 cycles" is an event measurable by the system.

Beside current events, the system can make judgments about past and future events, too. In NARS every sentence has a *time-stamp* indicating when the judgment is created (either from input or from inference); if the sentence is about an event, there is also a time-stamp about the estimated occurrence time. All the time-stamps are in terms of the system's internal clock, and each takes an integer as value, which can be either positive or negative. This treatment has some similarity with "step-logic" [5], though in NARS a time-stamp is not explicitly expressed as part of a statement. Unlike some cognitive architectures [17, 7], NARS does not attempt to simulate the response time of the human brain. The system uses its (subjective) working cycle as the unit of time, not the (objective) time provided by the clock of the host computer, so as to achieve platform-independence in testing. For example, if a certain inference process takes 10 steps in one computer, so does it in a different computer, even when the two systems have different running speeds.

The internal clock and built-in temporal relations are preferred for their simplicity and flexibility, but they are not used to represent all types of temporal information. NARS can use an external clock by specifying an event as occurring at the same moment as a time indicated by the clock. Since such a clock is an optional tool, the system can use different clocks in different situations for various demands of accuracy and granularity in time measurement.

In summary, in NARS temporal information is represented at three levels:

Term. A term (either atomic or compound) can represent a temporal concept (such as "New Year's Day") or relation (such as "after a while"). Such a term is handled just like the other terms, though its meaning contains acquired temporal information.

Statement. A temporal statement can be formed using a built-in *temporal* relation combined with certain *logical* connectors. For example, if A, B, and C are events, then the Narsese statement "$(A, B) \mathrel{/\!\!\!\Rightarrow} C$" represents "If A is followed by B, then C will occur after them".

Sentence. A temporal sentence uses a time-stamp to indicate the estimated occurrence time of the event, with respect to the internal clock of the system.

Since the internal clock is "private" to the system, when a temporal sentence needs to be expressed in Narsese for communication purpose, its time-stamp is converted into a "tense", which has three possible values: "past", "present", and "future", with respect to the "current moment" when the message is created. Symmetrically, when an input judgment has a tense attached, it is converted into a time-stamp, according to the current time.

It is important to see that an AGI system like NARS should not directly carry out inference on tense, because since the system works in real time, the "current moment" changes constantly [5, 12]. On this aspect, NARS is fundamentally different from many traditional temporal logic systems [22, 29], which treat the tense of a statement as one of its intrinsic properties, as if the reasoning system itself is outside the flow of time.

In summary, many different techniques have been proposed in AI to represent temporal information, each of which is effective under different assumptions [2]. NARS uses three approaches, and integrates them to satisfy the need of AGI.

2 Classical Conditioning as Temporal Inference

NARS uses *experience-grounded semantics* [30]. Accordingly, the truth-value of a statement measures its *evidential support* with two real numbers in [0, 1]: the *fre-*

quency value is the proportion of positive evidence among all currently available evidence, and the *confidence* value is the proportion of the currently available evidence among all evidence accumulated after the coming of new evidence by a unit amount. Their relation with probability is explained in [31].

Based on this semantics, each inference rule in NARS has a truth-value function calculating the truth-value of the conclusion according to the evidence provided by the premises. Without going into the details of the inference rules (covered in [32] and other publications on NARS), for the current discussion it is sufficient to know that as far as the *confidence* of the conclusion is concerned, there are three types of inference rules:

Strong Inference. For example, from premises "$\{Tweety\} \rightarrow bird\ \langle 1.00, 0.90\rangle$" ("Tweety is a bird") and "$(\$x \rightarrow bird) \Rightarrow (\$x \rightarrow [yellow])\ \langle 1.00, 0.90\rangle$" ("Birds are yellow", where $\$x$ can be substituted by another term), the *deduction* rule derives the conclusion "$\{Tweety\} \rightarrow [yellow]\ \langle 1.00, 0.81\rangle$" ("Tweety is yellow"). Such a rule is "strong" because the confidence of its conclusion can approach 1. If the truth-values are dropped and all the statements are taken to be "true", the rule is still valid in its binary form.

Weak Inference. For example, from premises "$\{Tweety\} \rightarrow bird\ \langle 1.00, 0.90\rangle$" and "$\{Tweety\} \rightarrow [yellow]\ \langle 1.00, 0.90\rangle$", the *induction* rule derives "$(\$x \rightarrow bird) \Rightarrow (\$x \rightarrow [yellow])\ \langle 1.00, 0.45\rangle$"; similarly, from "$(\$x \rightarrow bird) \Rightarrow (\$x \rightarrow [yellow])\ \langle 1.00, 0.90\rangle$" and "$\{Tweety\} \rightarrow [yellow]\ \langle 1.00, 0.90\rangle$", the *abduction* rule derives "$\{Tweety\} \rightarrow bird\ \langle 1.00, 0.45\rangle$". Such a rule is "weak" because the confidence of its conclusion cannot be higher than 0.5. If the truth-values are dropped and all the statements are taken to be "true", the rule becomes invalid in its binary form.

Evidence pooling. If two premises have the same statement but are supported by distinct evidence, such as "$bird \rightarrow [yellow]\ \langle 1.00, 0.50\rangle$" and "$bird \rightarrow [yellow]\ \langle 0.00, 0.80\rangle$", the *revision* rule derives "$bird \rightarrow [yellow]\ \langle 0.20, 0.83\rangle$". This is the only rule whose conclusion has a higher confidence value than both premises, since here the premises are based on the distinct evidential bases, while the conclusion is based on the pooled evidence.

There are many other inference rules in the system for other combinations of premises with respect to various term connectors, and they will not be addressed in this paper. In the following we only briefly describe how *temporal inference* is carried out. Here the basic idea is to process *temporal* information and *logical* information in parallel. Among other functions, this type of inference can carry out a process that is similar to *classical* (*Pavlovian*) *conditioning*, by associating a conditioned stimulus (CS) with an unconditioned stimulus (US). However what is special in NARS is that temporal inference will also happen between neutral stimuli. In the rare case that they get attention and also turn out to be important, they will find relations which a classical conditioning model would have missed.

To show how it works, assume initially the system gets to know that an occurrence of event C is followed by an occurrence of event U. As mentioned previously, events are represented as statements with temporal information. In this case, the occurrence time of C will be recognized by the system as before that of U. As soon as the temporal succession between the two events is noticed by the system, a temporal version of the *induction* rule will be invoked to generalize the observation into a temporal implication "$C \mathbin{/\!\!\!\Rightarrow} U$". The truth-value of this conclusion depends on the quality of the observations, as well as the restriction applied by the induction rule, so that the confidence value of the conclusion will be less than 0.5 – since it is only based on a single observation, the conclusion is considered a "hypothesis" that differs from a "fact" in confidence.

If at a later time C occurs again, then from it and the previous hypothesis the system derives U by deduction, with a time-stamp suggesting that it will occur soon. Since the hypothesis has a low confidence, the prediction on U is also tentative, though it may still be significant enough to raise the system's *anticipation* of the event, so as to make it more recognizable even when the input signal is relatively weak or noisy. An anticipation-driven observation is "active", rather than "passive" (where the system simply accepts all incoming signals without any bias), and the difference is not only in sensitivity. When expressed as Narsese sentences, the inputs provided by a sensor normally correspond to *affirmative* judgments, without any *negative* ones – we can directly see or hear what *is* out there, but cannot directly see or hear what *is not* there. "Negative observations" are actually *unrealized anticipations* and can only be produced by active observations.

In the current example, if the anticipated U does not appear at the estimated time, this unrealized anticipation and the preceding C will be taken as negative evidence by the *induction* rule to generate a negative judgment "$C \mathbin{/\!\!\!\Rightarrow} U$" that has a low (near 0) frequency value. Then the revision rule can pool this one with the previous (affirmative) one to get a new evaluation for the temporal statement "$C \mathbin{/\!\!\!\Rightarrow} U$". In this way, the successes and failures of anticipation will gradually lead the system to a relatively stable belief on whether, or how often, U is followed by C. The conclusion is similar to a statistical one, though it is revised incrementally, with no underlying probabilistic distribution assumed.

If the system has an unconditioned response (*UR*) to the *US*, this "instinct" corresponds to a temporal implication "$U \mathbin{/\!\!\!\Rightarrow} \Uparrow R$" that represents a sufficient precondition U for the operation $\Uparrow R$ to be executed, and it will have an affirmative truth-value, such as $\langle 1.00, 0.99 \rangle$ (confidence cannot reach 1, even for an instinct). From this instinct and the belief on "$C \mathbin{/\!\!\!\Rightarrow} U$", the *deduction* rule generates "$C \mathbin{/\!\!\!\Rightarrow} \Uparrow R$", which gives the operation an acquired sufficient precondition, though with a lower confidence than the instinct at the beginning. Now $\Uparrow R$ becomes a *conditioned response* (CR) to the CS.

Similarly, if the system already has a strong belief on "$C \mathbin{/\!\!\!\Rightarrow} U$", and it notices an occurrence of U, then by *temporal abduction* the system will guess that C has occurred previously, though the system may fail to notice it in the input stream, or it may be not directly observable. Similar to inductive conclusions, such an abductive conclusion is not very confident until it is strengthen by other evidence. As proposed by C. S. Peirce [20], a major function of abduction is to provide *explanations* for observations.

Most of the existing models of classical conditioning are built in the framework of dynamic system [24, 28, 8, 3], while in NARS it is modeled as an inference process. Though Bayesian models [27] also treat conditioning as reasoning, there the process only evaluates the probability of given statements, while NARS, following a logic, can generate new statements. Beside recognizing the preconditions and consequences of single operations, temporal inference also allows the system to do the same for *compound operations* consisting of multiple steps, which is usually called "planning", "scheduling", or "skill learning" [15]. Typically, the consequence of a preceding operation enables or triggers a following operation, and such a compound operation as a whole will gradually be used as an individual operation by the system. Such a process recursively forms an action hierarchy, which allows efficient reaction and planning with different granularity. Unlike in reinforcement learning or many other planning systems, NARS does not plan its actions in all situations in terms of the same set of basic operations.

3 Causal Inference without a Casual Relation

Based on the current situation to predict the future (or to describe the past) is often considered as "causal inference". Though "causality" has many different interpretations [33], a common opinion is to think the events in the universe as interconnected via "causal relations", so that every event is "caused" by a certain proceeding event. When the causal relations of an event become fully known, its occurrence and consequences can be predicted with certainty.

This opinion is usually formalized in AI and cognitive science using mathematical logic [9, 26], probability theory [4, 19, 23], or a combination of the two [11, 10]. Such a model assumes the existence of a deterministic or probabilistic *causal relation*, on which the system carries out logical or statistical inference to predict the future, or to describe the past, according to the present. In this approach, every event has a unique "true cause", which can be found, or at least approximated, using causal inference.[3]

NARS does not treat causal inference in this way. As mentioned previously, the basic assumption behind NARS is that an intelligent system never has full knowledge about the environment and itself, and all the available knowledge are revisable. For a given event, the system cannot know all of its preconditions so as to predict its occurrence with certainty. Similarly, the system cannot accurately anticipate all effects an event triggers. According to this opinion, even the probabilistic models assume too much – to meaningfully talk about "the probability" of an event, the presumption is that all the relevant events are in a space on which a probability function is defined. For an AGI system working in realistic situations, such a function can neither be obtained nor maintained, since the system does not know all future events, nor can it always guarantee the consistency among its degrees of belief when they are revised in real-time.

[3] A new approach [23] additionally tries to get rid of certain undesired results (Berkson's Paradox) of Bayesian conditioning by using "relational blocking", but the problem of assuming "true cause" remains.

In an AGI system, the above restrictions do not rule out the feasibility of predicting the future and describing the past. As shown by the previous example, NARS can learn the regularity in its experience, and use it to predict the future. Here the relevant knowledge is represented as temporal implication judgments like "$C /\!\!\!\Rightarrow U \ \langle f, c \rangle$", which is a summary of the relevant past experience, not an accurate or approximate description of an objective "law of nature".

The existence of objective causation is a long-lasting belief accepted by many scientists and philosophers, but it has been challenged in the recent century both in science (especially in physics) and in philosophy. From the point of view of cognitive science, it can be argued that all the beliefs of a system are restricted by the cognitive capability (*nature*) and past experience (*nurture*) of the system, and there is no ground to assume that such a belief will converge to an objective *truth*. Even so, an adaptive system can form beliefs about causation. According to Piaget [21], such beliefs originate from the observations about the consequences of one's own operations. For example, if event E is repeatedly observed after the execution of operation R, NARS will form a belief "$\Uparrow\!R /\!\!\!\Rightarrow E \ \langle 0.98, 0.99 \rangle$", which can be interpreted as "E is caused by R". This will be the case even when this achievement of the operation actually depends on a condition C, which is usually (say 98% of the time) satisfied – the belief is stable and useful enough for C to be ignored. However, when C is not usually satisfied, a belief like "$\Uparrow\!R /\!\!\!\Rightarrow E \ \langle 0.49, 0.99 \rangle$" will not be as useful to the system, so in this case a more reliable (though also more complicated) belief "$(C, \Uparrow\!R) /\!\!\!\Rightarrow E \ \langle 0.99, 0.95 \rangle$" will be favored by the system as the knowledge about how to get E. Please note that even in such a case it is hard to say what is the "true cause" for E to happen, since accurately speaking there may be other events involved, though for the system's current purpose, they do not need to be taken into consideration.

This discussion is also related to the Frame Problem [16], where the issue is: for a given operation of the system, how to represent all of its preconditions and consequences. The solutions proposed for this problem usually deal with it in idealized or simplified situations, while the response to it in NARS is to give up the attempt of getting all the information. An AGI system should depend on operations with incomplete descriptions of preconditions and consequences, and make decisions according to the available knowledge and resources [34].

NARS uses temporal inference to carry out prediction and explanation, which are often considered as "causal inference", though within the system there is no built-in "causal relation". The system has temporal versions of *implication* and *equivalence* relations built into its grammar and inference rules, so a "causal relation" can be represented in the system as their variant with domain-specific and context-dependent additional requirements. This treatment is arguably similar to the everyday usage of "causation". In many fields, questions of the form of "What is the real cause of X?", with various X, have been under debate for decades, even centuries. The notion of *cause* is interpreted very differently in different situations – it can be deterministic or probabilistic; it may correspond to a sufficient condition or a sufficient-and-necessary condition; it may or may not be an intentional action; and so on. However, behind all of these versions, the invariant components include a logical factor (from the given "causes", the "effects" can be derived) and a temporal factor (the "causes" happen no later than the "effects"). NARS covers these two aspects in temporal inference, while leaves the additional and variable aspects of causation to learning.

In this model, a *causal* relation and a *covariant* (or *correlative*) relation can still be distinguished, as usually desired [4]. However here their difference is *quantitative*, not *qualitative*. If the judgment on "$C /\!\!\!\Rightarrow U$" gets its truth-value solely by induction from a small amount of evidence, the confidence of the conclusion will be relatively low, and we tend to consider such a relation "covariant", but if the conclusion can also be established by a chain of deduction, such as from "$C /\!\!\!\Rightarrow M$" and "$M /\!\!\!\Rightarrow U$" where M is another event, then the relation between C and U may be considered as "casual", because it has an *explanation* leading to a high confidence. As far as prediction is concerned, what matters is the truth-value of the conclusion, not how they are derived. For instance, in Pavlovian conditioning the actual relation between CS and US is often coincidental, not causal, though animals in such experiments cannot tell the difference.

For a given event E, NARS can be asked to find its "cause" and "effect". The simplest form is to ask the system to instantiate the *query variable* $?x$ when answering questions "$?x /\!\!\!\Rightarrow E$" and "$E /\!\!\!\Rightarrow ?x$", respectively. When there are multiple candidate answers, a *choice* rule will be invoked to compare their truth-value, simplicity, relevance, etc., to pick the best answer. Additional requirements can be provided for the term or the statement that can be accepted as an answer. In general, NARS does not assume that such a question has a unique *correct* or *final* answer, but always reports the best answer it can find using the available knowledge and resources. Therefore, though the design of NARS does not include an innate causal relation, the system has the potential to *predict*, or even to *control*, the occurrence of an event. This is arguably what we should expect from an AGI.

4 Conclusions

Temporal inference plays a crucial role in AGI. An intelligent system needs the ability to learn the preconditions and consequences of each operation, to organize them into feasible plans or skills to reach complicated goals, and to find stable patterns among the events in its experience. This ability enables the system to predict the future, and to prepare sequences of operations to achieve its goals. Classical conditioning can be seen as a concrete case of this ability.

The approach of temporal inference in NARS allows temporal information to be expressed in several forms for different purposes. Some temporal notions are innate, while others are acquired, and they can be at different levels of granularity and accuracy. NARS integrates temporal inference with other inference, and utilizes a uniform memory for declarative, episodic, and procedural knowledge.

NARS carries out many cognitive functions, like prediction, that are usually associated with "causal inference". However there is no fixed notion of a "causal relation" within the system. NARS is based on the assumption that an accurate description of the universe with objective causal relations among the events may not be available to, or manageable by, the system, which makes NARS applicable to situations where many other models cannot be applied. Instead of trying to find or to approximate certain objective causal relations, what an intelligent system should do is to behave according to the regularity and invariance that it has summarized from its experience, and the generation, revision, and evaluation of such knowledge is a lifelong task.

All the aspects of NARS described in this paper have been implemented in the most recent version of the system. Currently the system, which is open source, is under testing and tuning. As an AGI system, NARS is not designed for any specific application, but as a testbed for a new theory about intelligence. Though the current implementation already shows many interesting and human-like properties, there are still many issues to be explored. This paper only addresses the aspects of NARS that are directly related to temporal inference.

Acknowledgment

The authors thank Kris Thórisson for helpful comments.

References

1. Allen, J.F.: Towards a general theory of action and time. Artificial Intelligence 23(2), 123–154 (1984)
2. Allen, J.F.: Time and time again: The many ways to represent time. International Journal of Intelligent Systems 6(4), 341–356 (1991)
3. Anderson, J.J., Bracis, C., Goodwin, R.A.: Pavlovian conditioning from a foraging perspective. In: Proceedings of the 32nd Annual Conference of the Cognitive Science Society. pp. 1276–1281 (2010)
4. Cheng, P.W.: From covariation to causation: a causal power theory. Psychological Review 104(2), 367–405 (1997)
5. Elgot-Drapkin, J., Perlis, D.: Reasoning situated in time I: Basic concepts. Journal of Experimental & Theoretical Artificial Intelligence 2, 75–98 (1990)
6. Franklin, S.: A foundational architecture for artificial general intelligence. In: Goertzel, B., Wang, P. (eds.) Advance of Artificial General Intelligence, pp. 36–54. IOS Press, Amsterdam (2007)
7. Franklin, S., Strain, S., McCall, R., Baars, B.: Conceptual commitments of the LIDA model of cognition. Journal of Artificial General Intelligence 4(2), 1–22 (2013)
8. Gallistel, C.R., Gibbon, J.: Time, rate, and conditioning. Psychological Review 107(2), 289–344 (2000)
9. Giunchiglia, E., Lee, J., Lifschitz, V., McCain, N., Turner, H.: Nonmonotonic causal theories. Artificial Intelligence 153, 49–104 (2004)
10. Goodman, N.D., Ullman, T.D., Tenenbaum, J.B.: Learning a theory of causality. Psychological review (2011)
11. Halpern, J.Y., Pearl, J.: Causes and explanations: A structural-model approach. Part I: Causes. The British journal for the philosophy of science 56(4), 843 (2005)

12. Ismail, H.O., Shapiro, S.C.: Two problems with reasoning and acting in time. In: Principles of Knowledge Representation and Reasoning: Proceedings of the Seventh International Conference. pp. 355–365 (2000)
13. Kowalski, R.: Logic for Problem Solving. North Holland, New York (1979)
14. Laird, J.E.: The Soar Cognitive Architecture. MIT Press, Cambridge, Massachusetts (2012)
15. van Lambalgen, M., Hamm, F.: The proper treatment of events. Blackwell Publishing, Malden, MA (2005)
16. McCarthy, J., Hayes, P.J.: Some philosophical problems from the standpoint of artificial intelligence. In: Meltzer, B., Michie, D. (eds.) Machine Intelligence 4, pp. 463–502. Edinburgh University Press, Edinburgh (1969)
17. Newell, A.: Unified Theories of Cognition. Harvard University Press, Cambridge, Massachusetts (1990)
18. Nivel, E., Thórisson, K.R., Steunebrink, B.R., Dindo, H., Pezzulo, G., Rodriguez, M., Hernandez, C., Ognibene, D., Schmidhuber, J., Sanz, R., Helgason, H.P., Chella, A., Jonsson, G.K.: Bounded recursive self-improvement. CoRR abs/1312.6764 (2013)
19. Pearl, J.: Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, UK (2000)
20. Peirce, C.S.: Collected Papers of Charles Sanders Peirce, vol. 2. Harvard University Press, Cambridge, Massachusetts (1931)
21. Piaget, J.: The construction of reality in the child. Basic Books, New York (1954)
22. Pratt-Hartmann, I.: Temporal prepositions and their logic. Artificial Intelligence 166, 1–36 (2005)
23. Rattigan, M.J., Maier, M., Jensen, D.: Relational blocking for causal discovery. In: Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (2011)
24. Rescorla, R., Wagner, A.: A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non reinforcement. In: Black, A., Prokasy, W. (eds.) Classical Conditioning II, pp. 64–99. Appleton-Century-Crofts, New York (1972)
25. Rosenbloom, P.S.: Rethinking cognitive architecture via graphical models. Cognitive Systems Research 12, 198–209 (2011)
26. Shoham, Y.: Nonmonotonic reasoning and causation. Cognitive Science 14, 213–252 (1990)
27. Srivastava, N., Schrater, P.: Classical conditioning via inference over observable situation contexts. In: Proceedings of the 36th Annual Meeting of the Cognitive Science Society. pp. 1503–1508 (2014)
28. Sutton, R.S., Barto, A.G.: Time-derivative models of Pavlovian reinforcement. In: M.Gabriel, Moore, J. (eds.) Learning and Computational Neuroscience: Foundations of Adaptive Networks, pp. 497–537. MIT Press (1990)
29. Vila, L.: A survey on temporal reasoning in artificial intelligence. AI Communications 7(1), 4–28 (1994)
30. Wang, P.: Experience-grounded semantics: a theory for intelligent systems. Cognitive Systems Research 6(4), 282–302 (2005)
31. Wang, P.: Rigid Flexibility: The Logic of Intelligence. Springer, Dordrecht (2006)
32. Wang, P.: Non-Axiomatic Logic: A Model of Intelligent Reasoning. World Scientific, Singapore (2013)
33. Williamson, J.: Causality. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. 14, pp. 95–126. Springer (2007)
34. Xu, Y., Wang, P.: The frame problem, the relevance problem, and a package solution to both. Synthese (2012)