

An Experience Is a Knowledge Representation

Keith McGreggor

College of Computing, Georgia Institute of Technology, Atlanta, GA, USA
keith.mcgreggor@gatech.edu

Abstract

Computational agents use knowledge representations to reason about the data world they occupy. A theory of consciousness, Integrated Information Theory, suggests beings that are conscious use experiences to reason about the world they occupy. Herein, the question is considered: is an experience a knowledge representation?

On Integrated Information Theory

Tononi has described and refined integrated information theory (IIT), a theoretical framework for describing and measuring consciousness (Tononi 2008, Tononi and Koch 2015). IIT argues that to construct a theory of consciousness, one must begin by creating a set of axioms based upon the phenomena to be explained, and then derive a set of postulates from those axioms. The five axioms of consciousness of IIT describe properties of an experience: existence, composition, information, integration, and exclusion (Tononi and Koch, 2015). The postulates IIT derives imply how the world that a conscious being occupies must be organized in order to support these properties of experience. A brief discussion of the postulates of IIT may be found here (Tononi 2015).

In IIT, the world that the conscious being experiences is said to be “real” only via derivation and reasoning from the experiences that that being has had. This directionality, that the world must be in some manner because the experience exists, is important, and the subject of discussion in the implications section.

What is an Experience?

Tononi et al. (2015) note that a proper theory of consciousness first must consider the essential properties of the phenomenon that the conscious being has had - its own “experience” of the phenomenon. IIT identifies five such properties, or axioms of consciousness.

An Experience Exists

The existence axiom holds that an experience has an existence that does not depend upon external observation. This is to say that the conscious being having the experience assigns a “truth” value to it. IIT argues that an experience in and of itself is the only fact available, that it has an intrinsic existence.

An Experience is a Composition

The composition axiom holds that an experience is a composition of phenomenological distinctions. These distinctions can be simple (e.g. “a triangle”) or structured (“a red triangle on the left”). Each of these distinctions within an experience can be noted separately.

An Experience Is Information

The information axiom holds that an experience is composed of a specific set of specific distinctions, and therefore differs from other experiences. This specificity combined with the ability to note the distinctions separately provides a means for judging the similarity of experiences.

An Experience is Integrated

The integration axiom holds that the specific composition of an experience may not be reduced into subsets of those phenomenological distinctions, as those subsets would become themselves specific experiences.

An Experience is Exclusive

The exclusion axiom stems from the IIT concept that consciousness is definite; that is, that experiences are specifically the set of distinctions they have, and that experiences occur at a consistent tempo and size. Therefore, an experience is specific set of distinct phenomenological distinctions which occur during a specific period in time: an experience excludes other experiences.

What is a Knowledge Representation?

Acts of cognition involve the manipulation of knowledge which is represented in some manner. While the term “representation” is quite commonplace and its usage familiar, Davis, Shrobe and Szolovits (Davis et al., 1993) specified what comprises a representation. Markman (1999) likewise offers a description of a representation. Both Markman (1999) and Nersessian (2008) provide insights into how to characterize representations.

Representations according to Davis

Davis et al. (1993) note that representations play five distinct, critical roles. Those roles are as a surrogate, as a set of ontological commitments, as a fragmentary theory of reasoning, as a medium for pragmatically efficient computation, and as a medium of human expression.

When a mind reasons about its world, this reasoning occurs internally, while the majority of what it reasons about exists externally. A representation then must act as a surrogate for things which exist outside the reasoning agency: direct interaction with real world objects are paralleled by operations upon the internal representations of those objects.

Davis et al. (1993) raise two significant points concerning surrogates: what is a surrogate a surrogate for, and what is the fidelity of a surrogate? Some correspondence between the surrogate and its counterpart in the world must be specified. With respect to fidelity, what attributes of the original are preserved, omitted, or implied with the surrogate must be addressed, for perfect fidelity is impossible. Representations, then, must be imperfect, and since reasoning operates upon representations, so to must reasoning itself arrive at imperfect conclusions, even if the reasoning process itself is sound.

Selecting a representation involves a decision about how and what to represent from the arriving world. A set of commitments, then, is made that both define the extent of the representation’s capture of the world and define the way that extent is expressed or embodied within the representation ontologically. Commitments start at the moment a representation begins to form, and accumulate as the representation is used. As Davis et al. (1993) note, the representational power lies in the correspondence of the representation to something in the world and in the constraints that that correspondence impose.

Representations are formed to allow cognition to occur within some agency. Even though the theory of reasoning arising from a representation may be implicit, it can be seen through three aspects: what the representation defines as inferencing, the set of inferences it allows, and the subset of those inferences which it recommends.

Allowed inferences are those inferences which can be made from available information. As a representation might arise in any number of ways, so too might the allowed inferences vary. As Davis et al. (1993) point out, this flexibility is acknowledged so as to admit the legitimacy of the various approaches. Having this flexibility at its core provides a framework for re-representation.

For a given representation, the set of allowable inferences may become untenably large. A smaller, constrained subset of these inferences is necessary. Whether by specifying the constraints with which to select recommended inferences, or by providing them somewhat explicitly, some process or reasoning or insight must be at work to frame them. In this way, Davis et al. (1993) citing Minsky by way of example, illustrates that representation and reasoning are intertwined in a deep, theoretical manner. They also observe that much of the reasoning which informs recommended inferences has been provided by observation of human behavior.

In the same sense that a representation recommends inferences, so to does it imply the manner in which it may be used in computation. This guidance speaks to the adequacy of the representation, as an organizational mechanism for information, for some task at hand.

Although Davis et al. (1993) addresses the notion of representations as vehicles for human expression, a computational agent’s internal dialogue of, about, and with representations is as important as an external one. In complex systems, information must pass from subsystem to subsystem, preferentially without substantial degradation and with increasing specificity. The expression of representations internally is a process of systematic reassembly of aspects of those representations into new ones, through which other systems may operate upon the newfound representations, with the core roles of representations implied by those systems’ tasks. Herein, cognitive models are formed. The text below a second-level heading begins without indentation. Use of the subsection heading style is required.

Representations according to Markman

In his book “Knowledge Representation,” Markman offers both a definition of representation as well as a set of criteria for assessing a representation (Markman, 1999). For Markman, a representation has four components: a represented world (the domain that the representation is about); a representing world (the domain which contains the representation); representing rules which map elements in the represented world to elements in the representing world; and a process which uses the representation.

Markman notes that in all known representational systems, the representing world loses information about the represented world, and assigns this loss of information to

the decision made about what aspects of the represented world to be included in the representing world. The agent constructing the representing world must decide what to include, and what to exclude, and that decision carries forward into the representation the consequences of it.

Markman notes that the representing rules determine the isomorphism (or homomorphism) of the representation: if each unique element in the represented world is mapped to a unique element in the representing world, the representation is isomorphic (Markman, 1999). The correspondence given by these representing rules also imply loss of information: if a representation is homomorphic, then more than one element in the represented world maps in an undifferentiable manner to the same element in the representing world, and therefore the ability to discriminate between those represented world elements is lost. This loss of information, through deliberate omission and through potential homomorphism, affords the capacity for reasoning about the missing information from that which is not missing.

Markman's association of a representation with some process which uses it implies that utility to an agent is the rationale for the construction of the representation. Markman additionally notes that Marr (1982) remarks that a given representation makes some information about the represented world easier to access than other information, via the representing rules and the loss of information.

Characterizations of Representations

There are additional ways to characterize representations, and both Markman (1999) and Nersessian (2008) provide insights into how to achieve such characterizations.

Analog / Symbolic

Markman distinguishes representations as either analog or symbol. A representation is an analog if the representing world has an inherent structure about how it operates and that the relationships between elements in the representing world are not arbitrary. A representation is symbolic if a convention exists which links all of the elements in the representing world, the convention being arbitrary in a sense that representing rules could be changed to determine a wholly new convention. The representing rules determine, therefore, the nature of a representation's analogism or symbolism.

Iconic / Propositional

Nersessian (2008) employs a slightly different terminology. To Nersessian, a representation may be characterized as as iconic if it demonstrates a structural relationship to the thing it represents. Iconic representations therefore afford an ability to assess similarity or goodness of fit, and provide a notion of being "accurate" or "inaccurate" (Nersessian 2008). Nersessian's iconic representation is closely associated with Markman's analog representation.

In contrast, Nersessian holds that if the relationship between a representation and what it represents stands for a kind of "truth" and if the operations over the representation preserve this "truth" via the use of a consistent set of symbols which themselves stand for a stable collection of properties, then the representation is propositional (Nersessian, 2008). Nersessian's propositional representation is closely aligned with Markman's symbolic representation.

Modal / Amodal

Nersessian (2008) further delineates representation along a dimension which pertains to the degree to which its symbols can be associated with perceptual states (Barsalou, 1999, 2008). Modal symbols are analog (in the Markman sense) representations of the perceptual states from which they are extracted. Amodal symbols, on the other hand, are arbitrarily (but consistently) assigned. Therefore, in Nersessian's view, a propositional representation uses amodal symbols, but an iconic representation may use either modal or amodal symbols, or both.

Assessment of Representations

Markman further suggests that proposed representations be assessed with respect to at least three dimensions: their endurance, the presence of symbols, and their abstractness. By endurance, Markman means not that some specific values within a representation be maintained (a state), but that the representation itself may be temporary or long-lasting. By the presence of symbols, this is a distinction between representations which are symbolic and which are not. Markman invokes the use of a space (as a structure upon which elements have some positional meaning) as an example of non-symbolic representation. Lastly, by abstractness, Markman suggests that this is the degree to which the process which uses the representation is distinct from the representation itself. Markman further develops the notion of the power of a representation as a convolution of another way in which to describe the suitability of the representation to the process which intends to use it with the expressivity of the representation (the degree to which it may be able to represent all represented worlds).

Experience, considered

Let us now consider whether experience may be a knowledge representation, working in reverse order, from Markman to Davis.

Experience and the worlds

An experience is an irreducible set of phenomenological distinctions which occur during a period of time. The represented world, though implicit, would be the world which gave rise to the phenomena. The representing world is a set

of specific experiences. There is some commitment made, via selection of phenomena, as to what aspects of the represented world are selected composed into the representing world. This commitment is the primary method by which the inclusion/omission of represented world information is made. An experience satisfies this aspect of Markman's definition.

Experience and the representing rules

Due to the selection of phenomena, some collection of representing rules must exist to distinctly transduce the represented world, at a definite tempo and granularity, into the representing world (set of experiences). Moreover, this mapping is likely isomorphic, to provide the irreducible nature of the experience. Therefore, an experience satisfies this aspect of Markman's definition.

Experience and symbolism

An experience is non-symbolic in the Markman sense, in that it rests upon an inherent structure given by the representing rules which is non-arbitrary. However, the phenomenological distinctions within an experience may themselves be represented according to any suitable arbitrary convention, so long as they allow for discrimination between themselves. While the symbols chosen may indicate correspondence to certain non-arbitrary aspects of the representing world (a position in space, a color, etc.), the manner in which the features are denoted itself is independent and arbitrary with respect to inherent structure of the experience and to the manner in which the comparison between features is made. Thus, phenomenological distinctions are symbolic in the Markman sense, but the experience of which they comprise is not. Even so, an experience satisfies this aspect of Markman's definition.

Experience and modality

An experience is specifically and clearly modal, for it expressly relates the phenomenological input from the represented world and establishes an isomorphic mapping between that input and the representing world.

Experience as a surrogate

Davis et al. (1993) argue that a knowledge representation is a surrogate for the world, over which reasoning is performed. An experience is an irreducible, definitive set of phenomenological distinctions. No reasoning about an experience involves the original, represented world: reasoning may only be performed on the set of experiences (or the distinctions that comprise them). An experience maintains a strong, direct correspondence between the represented and representing worlds, a consequence of the act of transducing the phenomena. Furthermore, the fidelity of the

correspondence is determined precisely by the transduction scheme. This commitment of an experience to correspondence and fidelity, driven largely by the transduction, allows the representation to satisfy the first role of Davis et al.

Experience as ontological commitments

An experience, through the transduction process by which it is derived from the represented world, clearly makes a deliberate commitment and mapping between the represented and representing worlds. But is this ontologically sound?

There is absolute grounding between each experience in the representing world and the phenomenological distinctions occurring during a period of time which transduced from the represented world. Moreover, the mapping is wholly isomorphic. There is no other potential meaning for any of the transductions other than they precisely stand as the capture of the mapping. Each experience is complete, concise, and deliberately excludes information from any other portion of the represented world. Thus, an experience satisfies the second role of Davis et al.

Experience as a fragmentary theory of reasoning

To consider experience as a fragmentary theory of reasoning, we must consider what the experience defines as inferencing, the set of inferences it allows, and the subset of those inferences which it recommends. Let us consider closely the afforded and sanctioned inferences of an experience.

An experience affords the ability to consider its composition of phenomenological distinctions. It also affords the consideration of each of those phenomenological distinctions. It would appear to admit a consideration of each experience as a member of a directed sequence, as it is transduced from a specific period of time. By allowing these considerations, it affords the ability to compare one experience to another.

An experience specifically sanctions all of the above as well, but it does not, perforce, sanction the partial combination of aspects of one experience with aspects of another. Why? Because to do so would negate the strong correspondence between the represented and representing worlds. Indeed, the only sanctioned operations are those which expressly maintains that correspondence. In this way, an experience satisfies the third role.

Experience as a medium for computation

The IIT consciousness theory supposes that experiences are the axiomatic foundation for reasoning about the reality of the world occupied by a conscious being. The ability to compare experiences and aspects of experiences without resorting to repeating the transduction of the represented world into experiences suggests that a set of experiences

is a computationally efficient computational substrate. In this manner, an experience satisfies the fourth role.

Experience as a medium for expression

This last role of Davis et al. proves the most vexing to argue for, for at its core, this would seem to require the communication of experiences between two conscious beings in order to assess its expressivity. However, as noted above, in complex systems information must pass from subsystem to subsystem.

An experience may need to be examined for its structure or distinctions in order for the conscious being to accomplish some task. The experience itself is far more compact than the original (and arguably infinite) data in the represented world. Thus, any subsystem which makes use of an experience or shares it with another (for example, from a memory system to a system which calculates similarity) benefits from the efficiency of this compaction.

The operation over experiences internally then may be considered a process of systematic reassembly of aspects of those experiences into new representations, over which other systems may operate, with the core roles of those new representations implied by those systems' tasks. Tononi and Koch (2015) note that through the several postulates derived from the axioms of consciousness, conceptual structures are formed. The expressivity of an experience grants these actions, and thus, an experience satisfies this fifth role.

An Experience is a Knowledge Representation

In light of the strength with which an experience satisfies the criteria of Markman and meets the roles of Davis et al., we claim that an experience is a knowledge representation. Furthermore, and to be quite specific ala Markman and Nersessian, an experience is an analog/iconic modal knowledge representation.

References

- Barsalou, Lawrence W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, 22(04), 577-660.
- Barsalou, Lawrence W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59, 617-645.
- Davis, Randall, Shrobe, Howard, & Szolovits, Peter. (1993). What is a knowledge representation? *AI magazine*, 14(1), 17.
- Markman, Arthur B. (1999). *Knowledge representation*: Routledge.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: WH Freeman and Co.
- Nersessian, Nancy. (2008). *Creating scientific concepts*: The MIT Press.

Tononi G (2008). Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, 215 (3), pp. 216–242.

Tononi G (2015). Integrated information theory. *Scholarpedia*, 10 (1), pp. 4164.

Tononi G, Boly M, Massimini M, Koch C (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17 (7), pp. 450–461.

Tononi G, Koch C (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370 (1668), pp. 20140167.