

A Layered Abduction Model of Perception: Integrating Bottom-up and Top-down Processing in a Multi-Sense Agent

John R. Josephson
Laboratory for Artificial Intelligence Research
Ohio State University
Columbus, Ohio 43210

1 February 1989

Abstract

This paper introduces a layered-abduction model of perception which unifies bottom-up and top-down processing in a single logical and information-processing framework. The process of interpreting the input from each sense is broken down into discrete layers of interpretation, where at each layer a "best explanation" hypothesis is formed of the data presented by the layer or layers below, with the help of information available laterally and from above. The formation of this hypothesis is treated as a problem of abductive inference, similar to diagnosis and theory formation. Thus this model brings a knowledge-based problem-solving approach to the analysis of perception, treating perception as a kind of "compiled" cognition.

The bottom-up passing of information from layer to layer defines channels of information flow, which separate and converge in a specific way for any specific sense modality. Multi-modal perception occurs where channels converge from more than one sense.

This model has not yet been implemented, though it is based on systems which have been successful in medical and mechanical diagnosis and medical test interpretation.

Introduction

Computational models of information processing for both vision and spoken language recognition have commonly supposed an orderly progression of layers, beginning near the retina or auditory periphery, where hypotheses are formed about "low-level" features, e.g., edges (in vision) or bursts (in speech perception), and proceeding by stages to higher-level hypotheses. These higher-level hypotheses typically depend largely on hypotheses formed at lower levels, but are also subject to influence from above.

Models intended to be comprehensive often suppose 3 or more major layers, often with sublayers, and sometimes with parallel channels which separate and combine to support higher-layer hypotheses (e.g., shading discontinuities and color contrasts separately supporting hypotheses about object boundary) [31, 28, 29]. Audition, Phonetics, grammar, and semantics have been proposed as layers of interpretation for speech comprehension. Recent work on primate vision appears to show the existence of separate channels for information about shading, texture, and color, not all supplying information to the same layers of interpretation [29].

In both vision and speech understanding most of the processing of information is presumably bottom-up, from information produced by the sensory organ, through intermediate representations, to the abstract cognitive categories that are required for reasoning. Yet top-down processing is significant, as higher-level information is brought to bear to help with identification and disambiguation. Both vision and speech recognition can thus be thought of as "layered interpretation" tasks whereby the output from one layer becomes data to be "interpreted" at the next. Layered interpretation models for non-perceptual interpretive process make

sense too, for example medical diagnosis can be thought of as an inference which proceeds from symptoms, to pathophysiological states, to diseases, to etiologies. It is reasonable to expect that perceptual processes have been optimized over evolutionary time, and that the specific layers and hypotheses, especially at lower levels, have been compiled into special-purpose mechanisms. Perceptual learning provides another source of compilation and optimization. Nevertheless, these layered interpretation models all seem to share certain functional similarities.

In particular, it appears that at each layer of interpretation the information processing task is the same: that of forming a coherent, composite (multi-part) “best explanation” of the data from the previous layer. That is, the task is one of performing an *inference to the best explanation*, in other words, an *abductive inference*.

Moreover, it appears that similar types of hypotheses-hypothesis interactions appear in vision, speech understanding, and diagnosis. Here are three important ones:

1. Two hypotheses might partially overlap in what they can account for, but otherwise be compatible (e.g. an edge might be a boundary for two different objects, the /s/ sound acoustically in the middle of “six stones” belongs to both words, the high white blood count is a result of two different infections),
2. Hypotheses might be pair-wise incompatible (e.g. patch X is either part of the figure or part of the background),
3. Hypotheses might be supportive in an associative way, the presence of one giving some evidence for the presence of the other. (Associative support presumably represents the net impact of several distinct types of evidential relationships.)

The functional similarities suggests the possibility of a generic mechanism, and just such a generic mechanism is proposed here. It is hypothesized that that the processing that occurs in vision, hearing, understanding spoken language, and in interpreting information from other senses (natural and robotic) can all be usefully thought of as variations, incomplete realizations, or compilations (domain-specific optimizations) of this one basic computational mechanism, which we may call *the layered abduction model of perception*.

There is a long tradition of belief that perception involves some form of inference [27] [17] [2]. Several researchers have in fact proposed that perception, or at least language understanding, involves some form of abduction or best-explanation inference [10, p.557] [9] [11] [37] [21, pp.87-94] [14, pp.88,104]. Abduction is often thought of as being logically similar to theory formation in science [17] [46] [14, p.104] and to diagnostic reasoning.

Abduction

The logician and philosopher Charles Sanders Peirce introduced the term “abduction” to refer to a kind of plausible inference, which he took to be logically distinct from both induction and deduction [36]. An abduction passes from a body of data, to a hypothesis that *explains* or *accounts for* that data. Thus abduction is a kind of theory-forming or interpretive inference. In fact Peirce says in one place, “Abductive inference shades into perceptual judgment without any sharp line of demarcation between them.” [37, p.304].

In their popular AI textbook Charniak and McDermott characterize abduction variously as *modus ponens* turned backwards, inferring the causal reasons behind something, generation of explanations for what we see around us, and inference to the best explanation [10]. They write that medical diagnosis, story understanding, vision, and understanding natural language are all abductive processes, and they speculate as to whether there might be possible a “‘unified theory’ of abduction” which will link all of these processes together [10, p.557].

Other AI practitioners have given similar characterization of abduction [26] [38] [39] [40], some have proposed or built systems using similar ideas without actually using the term “abduction” in describing their work [3] [33] [12] [35] [41] [34]. Some attempts have been made to cast the natural language understanding problem explicitly as abduction [11] [9]. Philosophers have written of “inference to the best explanation” [19]

[13] [21] and “the explanatory inference” [30]. Related philosophical traditions are the “hypothetico-deductive” model of the scientific method, and accounts of “the logic of discovery” [18]. Recently Paul Thagard has given abduction an important role in his analysis of the logic of scientific theory formation [46].

We may characterize **Abduction** as a form of inference that follows a pattern like this:

D is a collection of data (facts, observations, givens),
H explains D (would, if true, explain D),
No other hypothesis explains D as well as H does.

Therefore, H is probably true.

The confidence in the conclusion should (and typically does) depend on these factors:

- how decisively H surpasses the alternatives,
- how good H is by itself, independently of considering the alternatives (e.g. we will be cautious about accepting a hypothesis, even if it is clearly the best one we have, if it is not sufficiently plausible in itself),
- how thorough the search was for alternative explanations, and
- pragmatic considerations, including
 - the costs of being wrong and the benefits of being right,
 - how strong the need is to come to a conclusion at all, especially considering the possibility of seeking further evidence before deciding.

I hope my reader recognizes this form of inference as being common in ordinary life, and a part of the “scientific method”. What I am proposing here is that it also occurs on many levels in perception.

In general, as Marr pointed out, it is important to distinguish the *goal* of a computation; from the *logic of the strategy* by which that goal can be achieved, from the specific *representations and algorithms* used to describe a specific strategy, and from *implementations* of those representations and algorithms [31, p.25]. Describing a layer of interpretive inference as “abduction” describes the goals of the inference, and suggests strategies to achieve them, as I hope will become clear in what follows. The discussion here will not directly address representation, algorithm, or implementation.

The Layered Abduction Model

Each layer of interpretation, or more precisely, each locus of hypothesis formation (leaving open the possibility of more than one per layer) I call an **agora** after the meeting place where the ancient Greeks would gather for dialog and debate. The picture is that an agora is a place where hypotheses of a certain type gather and contend and where under good conditions a consensus hypothesis emerges. In typical cases the emerging interpretive hypothesis will be a composite hypothesis, coherent in itself, and with different sub-hypotheses accounting for different portions of the data. For example in vision the edge agora can be thought of as the location where a set of edge hypotheses are formed and accepted, each specific edge hypothesis accounting for certain specific data from lower-level agoras.

Our model calls for the information processing at each agora to be decomposed into three functionally distinct types of activity, which we can call *evocation of hypotheses*, *instantiation of hypotheses*, and *composition of hypotheses*.

Evocation can occur bottom-up, a hypothesis being stimulated for consideration by the data presented at the layer below. In diagnosis we would say that the presence of a certain finding *suggests* that certain

hypotheses are appropriate to consider. More than one hypothesis may be suggested by a given datum. Evocation can also occur top-down, either as the result of priming (an expectation from the level above), or as a consequence of data-seeking activity from above, which can arise from the need for evaluation. Evocations can in general be performed in parallel, and need not be synchronized.

Instantiation occurs when each stimulated hypothesis is independently scored for confidence (*evaluation*), and a determination is made of what part or aspect of the data the hypothesis can account for (*determination of explanatory scope*). This process is in general top-down, and in order to instantiate itself a hypothesis may seek data which was not part of its original stimulus¹. The data which are accounted for may or may not be identical to the data upon which the hypothesis was scored, or the data which did the evoking.

In the course of instantiation the hypothesis set may be expanded by including subtypes and supertypes of high-confidence hypotheses². Instantiation is typically based on matching against prestored patterns of features, but instantiating “by synthesis” is also possible whereby the features to match are generated at run time. The result of a wave of hypothesis instantiation is a set of hypotheses, each with some measure of confidence, and each offering to account for some portion of the data. Usually many of the evoked hypotheses can be ruled out, and will not form part of the result. Since in a wave of instantiation hypotheses are considered independently of each other, this too can go on in parallel.

Composition occurs when the instantiated hypotheses interact with each other and (under good conditions) a coherent best interpretation emerges. Note that, as this stage begins, each hypothesis has both a confidence value, and a body of data that it can account for. In the end some hypotheses will have been incorporated into the composite hypotheses, some will have been excluded, and perhaps some will be in limbo as a result of some remaining ambiguity of interpretation.

Strategy for Composition of Hypotheses

For hypothesis composition an overall abduction problem has been set up: to account for all of the (reliable and important) data presented by the agora(s) immediately below. A series of small abduction problems is also set up: to account for each particular datum. A basic strategy is to try to solve the overall abduction problem by solving a sufficient number of smaller and easier abduction problems. We begin by solving the easiest small abduction problems, the ones in which we can have the most confidence. If a certain hypothesis is *the only plausible explanation* for some finding (it accounts for the finding and its local-match confidence value is not too low), then it is entitled to high confidence, and entitled to be accepted into the overall composite hypothesis that represents the solution to the overall abductive problem.

Let us call a hypothesis “BELIEVED” when it has been accepted as the correct interpretation for the data it offers to account for. Data accounted for by BELIEVED hypotheses are “ACCOUNTED-FOR” and are considered to be successfully interpreted. Let us call a hypothesis “ESSENTIAL” if it is the only plausible explanation for some reliable datum (which is typically a hypothesis at the next lowest level that is BELIEVED). Thus an ESSENTIAL hypothesis scores positively and accounts for data items for which there are no other good interpretations. ESSENTIAL hypotheses are BELIEVED. Information about the explanatory relationships is thus used to increase the confidence in certain hypotheses.

If not all of the data are yet accounted for, the next step is to propagate the consequences of the initial set of BELIEVED hypotheses. These consequences arise as the result of causal and statistical relationships between hypotheses typically stored as compiled knowledge in advance of processing. There are several kinds of these relationships—I describe them here just briefly. Hypotheses at the same level (in the same agora) can have relationships of compatibility, entailment, or incompatibility, which can be a matter of degree. Propagating the consequences of BELIEVED hypotheses by taking account of these relations requires the appropriate adjustment of scores for related viable hypotheses outside of the BELIEVED set, or other appropriate actions. For example a hypotheses incompatible with a BELIEVED hypotheses can be rejected categorically, and removed from further consideration. Another kind of relationship is where a hypotheses “EXPECTS” the

¹ Under certain data-driven circumstances it is good enough just to score on the basis of voting by the stimulating data from below, and then no top-down processing need occur, at least for scoring.

² In general the space of potential hypotheses can be assumed to be hierarchically organized by level of specificity.

presence of data items at the next lower level (this too can come in degrees). Propagating such an expectation requires evoking the hypothesis corresponding to the expectation (priming) if it has not already been evoked. If it has, it can be given an extra measure of confidence. If a strong expectation is *contradicted* by the data, an anomaly has occurred, and special handling is appropriate.³

A hypothesis is a "CLEAR-BEST" if it is the distinctly best explanation (by confidence level) of some data item. CLEAR-BEST hypotheses are BELIEVED too. Note that an ESSENTIAL or CLEAR-BEST hypothesis is the uniquely best explanation for some data items—it is a local abductive conclusion. If not all of the data has been accounted for, the consequences of CLEAR-BESTs are propagated similarly to the ESSENTIAL hypotheses. Note that this propagation can result in more hypotheses becoming CLEAR-BESTs, e.g., if high-scoring explanatory competitors are removed from consideration, or if propagating consequences readjusts hypothesis scores so that a clear winner emerges.

If the ESSENTIAL hypotheses together with the CLEAR-BESTs do not account for everything, we have done all we can do on the current evidence without resorting to guessing. Generally our best strategy under these circumstances would be to go back for more data. In fact we are in a position to guide the data gathering by focusing on the problem of discriminating between alternative good explanations for significant data items. This is a form of top-down processing we may call "focused disambiguation". Sometimes, however, we have all of the relevant data we are going to get, for example we may be unable to ask the speaker to repeat. Under these circumstances we still have the means available to do some clever guessing. We can begin to include hypotheses which are best explanations for certain findings, but which are not far enough ahead of the alternatives, or not of high enough local-match confidence, to enable them to be accepted confidently. These WEAKLY-BESTs constitute the best guesses we can make under the circumstances. Actually some of them can be accepted with a fairly high degree of confidence. A finding can be made to vote for the hypotheses which best explain it (with voting strength in proportion to the measure by which the hypothesis beats its nearest competitor). The idea is that two different findings, both pointing to the same hypotheses as the best explanation constitute (apparently) independent sources of evidence for the hypothesis, i.e., constitute converging lines of inference for the hypothesis. Hypotheses with more votes can be accepted more confidently than hypotheses with fewer votes, and perhaps enough can be confidently accepted to complete the explanation.

Now in general relationships (spatial, grammatical, etc.) between the parts of a hypothesis are significant and need to be maintained. Some of these relationships can be seen as the filling of related roles in higher-layer interpretive hypotheses, for example a diagnostic hypothesis of a flow going on between A and B would bind A-related and B-related data together into relationships. But some other relationships (e.g. spatial in low level vision) are presumably compiled into the hardware, so that the appropriate constraints are applied between neighboring hypotheses as an automatic result of the operation of the machinery. Still and all, the net impact on hypothesis composition of these relationships can probably be captured by basic relationships of mutual sympathy and antipathy.

At the end of a wave of composition activity certain hypotheses have been accepted as BELIEVED. These constitute a confident best explanation for a portion of the data. Often there will also remain a set of unexplained data, and a set of viable hypotheses which, at various levels of confidence, offer to explain that data, but for which no clear solution is apparent. Nevertheless the BELIEVED hypotheses may be enough data for the next higher layer to do its business; resolving the remaining ambiguities may be unimportant in the context. Alternatively, remaining ambiguities may get resolved later as a result of further processing at that layer stimulated by downward-flowing expectations.

Downward-Flowing Processing

We may distinguish at least four sources or functions of top-down processing. One is that the data-seeking needs of hypothesis evaluation can provoke computation of the data (top-down evocation and evaluation of a hypothesis) as was discussed above. Another that was mentioned is that expectations based on firmly established hypotheses at one layer can prime certain data items (i.e. evoke consideration of them and bias

³Throughout the processing various kinds of anomalies can occur. Anomalies are detected and recorded, and typically stimulate special handling; from here on I describe the course of processing only for when everything goes smoothly.

their score upwards). A third way is that hypotheses that are uninterpretable as data at the higher level (no explanation can be found) can be “doubted” and reconsideration of them provoked. Finally data pairs that are jointly uninterpretable, as for example two words, the co-occurrence of which cannot be reconciled syntactically or semantically, can be considered to be incompatible (to some degree of strength) and recomputation of the composite hypothesis can be provoked from above. In these ways higher-level interpretations can exert a strong influence on the formation of hypotheses at lower levels, and layer-layer harmony is a two way street.

Recovering from Mistakes

Mistakes in Initial Hypothesization and Scoring

- Hypothesis suggestions come from above as well as below, thus hypotheses which would be missed on bottom-up processing can still be considered.
- If suggestions are inadequate, e.g. no hypotheses are evoked covering a segment of data, or all suggestions score low, exhaustive search (though hierarchically organized for efficiency) is undertaken to broaden the hypotheses being considered, thus hypotheses that are missed on suggestion-based stimulation can still be considered.
- Hypothesis evaluation is augmented by encouragement and discouragement (resulting from positive associations and incompatibilities) from other hypotheses in the same agora. Thus the local-match confidence score is improved by contextual information.
- Hypotheses evaluation is augmented by encouragement and discouragement based on expectations derived from confident higher-level hypotheses. This constitutes another kind of context-based improvement and check on the the confidence score.
- The acceptance of a hypothesis is based on how well it surpasses explanatory alternatives, thus after recognition-based scoring, a significant additional uncertainty-reducing operation is performed before acceptance.
- Strength of confidence is supported by “the consilience of inductions” whereby converging lines of inference all support the same hypotheses. Thus system performance should be robust.
- Acceptance, when it finally occurs, is still tentative and liable to be overthrown by relationships to the mass of other confident hypotheses.

Mistakes in Choice of Initial Islands of Confidence

- Actually the islands are very strong. They are never based only on a hypothesis having high initial confidence; it is at least required to also be a distinctly best explanation for some datum.
- Inconsistencies lead to detected anomalies, which lead to special strategies that weigh alternative courses of action. Originally accepted hypotheses can collide with others and subsequently called into question.
- Inconsistency collisions can occur laterally, or from above (violation of expectation, or from below (violation of expectation), and can come in degrees of strength. In effect there is broad cross checking of accepted hypotheses.
- An inexplicable datum should be doubted and called into question—it may not really be there. If after re-evaluation the datum remains strong despite the doubt, then the system can detect that it has encountered the limits of its knowledge, and is positioned to learn a new hypothesis category.
- Sometimes two parts of a compound hypothesis are inconsistent in context, where the judgment of higher levels is that they cannot both occur, based upon the inability to form a consistent hypothesis at the next highest level. (It seems that this can account for unstable perceptual objects like the Necker cube.)

Summary of the Control Strategy

We may summarize the control strategy by saying that it employs multi-level and multiple intra-level island-driven processing. Islands of relative certainty are seeded by local abductions and propagate laterally (incompatibilities, positive associations), downwards (expectations), and upwards (firm items of data to be accounted for). Processing occurs concurrently and in a distributed fashion. Higher levels provide *soft* constraints through the impact of expectations on hypothesis evocation and scoring, but do not strictly limit the hypothesis space.

Extension of the Model to Multi-Modal Perception

The basic idea in extending the model to multi-modal perception, i.e. perception that combines the information from more than a single sense, is that combining information from different senses is functionally no different than combining information from different channels within one sense modality. Different channels within the visual system deliver up the data useful at a certain level to form hypotheses about the locations of 3-d objects within the visual space; similarly, different senses deliver up the data useful for forming hypotheses about, say, object identity.

One special processing problem for multi-sense integration is the problem of identifying a "That" delivered up by one sense, with a "That" delivered up by another. Which person is the one that is speaking? Is it the same object being seen in the infrared as that being seen in x-rays? Logically, it should be possible for information derived from one sense to help with resolving distinct objects within the other sense. There is actually some evidence that vision can help hearing to separate distinct streams of tones [32, p.83] and hear the tone stream as two distinct auditory objects.

One useful computational support for cross modal perception is provided by correlated spatial representations, as our visual maps are correlated with our auditory maps of the space surrounding us. Thus, for example, a robot should bring together separate channels of information from its senses of "sight" and "touch" into a unified spatial representation of its immediate surroundings. Moreover this "hot map" of its surroundings should be maintained continually, and updated and revised as new information arrives and is interpreted. This hot map, with its symbols on it, can be viewed as the resulting composite hypothesis formed at "the agora of objects in the immediate surroundings" by a process of abductive interpretation.

Yet some senses are not particularly spatial (e.g. smell). We can envision computational support for cross-modal perception in the form of pattern-based recognition knowledge, where the compiled recognition patterns for an object category rely on features from more than one sense. This is very analogous to medical diagnosis where a disease is recognized from evidence from such disparate sources as lab tests, x-rays, and patient history. Such recognition knowledge can be used to support an "agora of the patient's disease" in much the same manner as the robot mentioned above maintains its map of objects in its surroundings. Somewhat further along we can envision a robot that maintains an "agora of understanding" whereby it monitors some complex device and continually maintains a causal understanding of it. Much much further along we can imagine building a robot scientist who maintains an "agora of theoretical understanding" whereby its best understanding of the world is maintained.

Summary: Perception as Compiled Cognition

The formation of a composite best-explanation hypothesis at any level in perception is treated as a problem of abductive inference, similar to diagnosis and theory formation. Thus this model brings a knowledge-based problem-solving approach to the analysis of perception, treating perception as a kind of "compiled" cognition.

Acknowledgments

The model described here is an outgrowth of research on diagnostic reasoning that has gone on at Ohio State Laboratory for AI Research over the past decade in the context of a developing theory of generic information

processing tasks [4], [6], [5], [7], [8], [42], [43], [44]. Various facets of abductive reasoning have been investigated, especially the problem of assembling composite explanatory hypotheses [24], [25], [26], [1], [20], [16], [15], [45], [23] [22].

This work has been supported by the Defense Advanced Research Projects Agency under RADC contract F30602-85-C-0010; and the NIH Heart, Lung and Blood institute, grant number 1 RO1 HL38776-01. The development of the layered abduction model has benefited especially from discussions with B. Chandrasekaran, Susan Josephson, Dean Allemang, Ashok Goel, Tom Bylander, Michael Tanner, Terry Patten, and Jordan Pollack, as well many discussions at the Image Understanding Workshop sponsored by DARPA/ISTO in April 1988.

References

- [1] D. Allemang, M. C. Tanner, T. Bylander, and J. R. Josephson. On the computational complexity of hypothesis assembly. In *Proc. Tenth International Joint Conference on Artificial Intelligence*, Milan, August 1987.
- [2] Jerome S. Bruner. On perceptual readiness. *Psychological Review*, 64(2):123–152, 1957.
- [3] B. G. Buchanan, Sutherland, G. L., and E. A. Feigenbaum. *Heuristic DENDRAL: a program for generating explanatory hypotheses in organic chemistry*, pages 54–69. Edinburgh University Press, Edinburgh, 1969.
- [4] B. Chandrasekaran. Generic tasks in knowledge-based reasoning: High-level building blocks for expert system design. *IEEE Expert*, pages 23–30, 1986. Fall 1986.
- [5] B. Chandrasekaran. Towards a functional architecture for intelligence based on generic information processing tasks. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, 1987.
- [6] B. Chandrasekaran. Generic tasks as building blocks for knowledge-based systems: The diagnosis and routine design examples. Technical report, The Ohio State University, 1988.
- [7] B. Chandrasekaran, F. Gomez, S. Mittal, and J. W. Smith. An approach to medical diagnosis based on conceptual structures. In *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, pages 134–142, 1979.
- [8] B. Chandrasekaran and S. Mittal. Conceptual representation of medical knowledge for diagnosis by computer: Mdx and related systems. In M. Yovits, editor, *Advances in Computers*, volume 22, pages 217–293. Academic Press, 1983.
- [9] E. Charniak. A neat theory of marker passing. In *Proceedings of AAAI-86*, volume 1, pages 584–588. AAAI, Morgan Kaufmann, August 1986.
- [10] Eugene Charniak and Drew McDermott. *Introduction to Artificial Intelligence*. Addison Wesley, 1985.
- [11] Venugopala Rao Dasigi. *Word Sense Disambiguation in Descriptive Text Interpretation: A Dual-Route Parsimonious Covering Model*. PhD thesis, University of Maryland, College Park, 1988.
- [12] Johan de Kleer. Reasoning about multiple faults. *AI Magazine*, 7(3):132–139, August 1986.
- [13] R. Ennis. Enumerative induction and best explanation. *The Journal of Philosophy*, LXV(18):523–529, September 1968.
- [14] Jerry Fodor. *The Modularity of Mind*. Bradford Book, 1983.
- [15] Ashok Goel, J. Ramanujan, and P. Sadayappan. Towards a ‘neural’ architecture for abductive reasoning. In *Proceedings of the Second International Conference on Neural Networks*, volume 1, pages 681–688, 1988.

- [16] Ashok Goel, P. Sadayappan, and John R. Josephson. Concurrent synthesis of composite explanatory hypotheses. In *Proceedings of the Seventeenth International Conference on Parallel Processing*, pages 156–160, August 1988.
- [17] Richard L. Gregory. Perception as hypotheses. In Richard L. Gregory, editor, *The Oxford Companion to the Mind*, pages 608–611. Oxford University Press, 1987.
- [18] N.R. Hanson. *Patterns of Discovery*. Cambridge University Press, 1958.
- [19] Gilbert Harman. The inference to the best explanation. *Philosophical Review*, LXXIV:88–95, January 1965.
- [20] John Josephson. A framework for situation assessment: Using best-explanation reasoning to infer plans from behavior. In *Proceedings of Expert Systems Workshop*, pages 76–85, April 1987.
- [21] John R. Josephson. *Explanation and Induction*. PhD thesis, The Ohio State University, 1982.
- [22] John R. Josephson. Reducing uncertainty by using explanatory relationships. In *Proceedings of the Space Operations Automation and Robotics Conference-1988 (Soar-88)*, pages 149–151, 1988.
- [23] John R. Josephson. Towards a generic architecture for layered interpretation tasks—implications for plan recognition. Technical report, The Ohio State University, 1988.
- [24] John R. Josephson, B. Chandrasekaran, and Jack Smith, Jr. Assembling the best explanation. In *Proceedings of the IEEE Workshop on Principles of Knowledge-Based Systems*, pages 185–190, Denver, Colorado, December 1984. IEEE Computer Society. A revised version by the same title is available as a technical report.
- [25] John R. Josephson, B. Chandrasekaran, Jack Smith, Jr., and Michael C. Tanner. Abduction by classification and assembly. In *PSA 1986 Volume One*, volume 1, pages 458–470, East Lansing, Michigan, 1986. Philosophy of Science Association.
- [26] John R. Josephson, B. Chandrasekaran, Jack Smith, Jr., and Michael C. Tanner. A mechanism for forming composite explanatory hypotheses. *IEEE Transactions on Systems, Man and Cybernetics, Special Issue on Causal and Strategic Aspects of Diagnostic Reasoning*, SMC-17(3):445–54, May, June 1987.
- [27] Immanuel Kant. *Critique of Pure Reason*. St. Martins Press, (1968), New York, 1787. tr. Norman Kemp Smith.
- [28] V.R. Lesser, R.D. Fennell, L.D. Erman, and R.D. Reddy. The Hearsay II speech understanding system. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-23:11–24, 1975.
- [29] Margaret Livingstone and David Hubel. Segregation of form color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, May 1988.
- [30] William G. Lycan. Epistemic value. *Synthese*, 64:137–164, 1985.
- [31] David Marr. *Vision*. W. H. Freeman and Company, 1982.
- [32] Dominic William Massaro. *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry*. Lawrence Erlbaum Associates, New Jersey, 1987.
- [33] Randolph A. Miller, Harry E. Pople, Jr., and Jack D. Myers. Internist - i, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, 307:468–476, 1982.
- [34] R. S. Patil. *Causal representation of patient illness for electrolyte and acid-base diagnosis*. Ph. D Thesis, 1981.
- [35] J. Pearl. Distributed revision of composite beliefs. *Artificial Intelligence*, 33(2):173–215, October 1987.
- [36] Charles S. Peirce. Abduction and induction. In Justus Buchler, editor, *Philosophical Writings of Peirce*, chapter 11, pages 150–156. Dover, 1955.

- [37] Charles S. Peirce. Perceptual judgments. In Justus Buchler, editor, *Philosophical Writings of Peirce*, pages 302–305. Dover, 1955.
- [38] H. Pople. On the mechanization of abductive logic. *Proceedings of the Third International Joint Conference on Artificial Intelligence*, pages 147–152, 1973.
- [39] James Reggia. Abductive inference. In Kamal N. Karna, editor, *Proceedings of The Expert Systems in Government Symposium*, pages 484–489. IEEE Computer Society Press, 1985.
- [40] James A. Reggia, Barry T. Perricone, Dana S. Nau, and Yun Peng. Answer justification in diagnostic expert systems—part 1: Abductive inference and its justification. *IEEE Transactions on Biomedical Engineering*, BME-32(4):263–267, April 1985.
- [41] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32(1):57–95, 1987.
- [42] Jack Smith, Jr., John R. Svirbely, Charles A. Evans, Pat Strohm, John R. Josephson, and Michael C. Tanner. Red: A red-cell antibody identification expert module. *Journal of Medical Systems*, 9(3):121–138, 1985.
- [43] Jack W. Smith, Jr. *RED: A Classificatory and Abductive Expert System*. PhD thesis, Ohio State University, August 1985.
- [44] Jon Sticklen. *MDX-2: An Integrated Medical Diagnostic System*. PhD thesis, The Ohio State University, 1987.
- [45] Mike Tanner and John Josephson. Abductive justification. Technical report, Ohio State University Laboratory for AI Research, 1988.
- [46] Paul Thagard. *Computational Philosophy of Science*. Bradford Books / MIT Press, 1988.