duced by the other participant. She argued that this ingrained conversational collaboration should be exploited to design successful natural language interfaces.

Paul McKevitt, New Mexico State University, described a Wizard of Oz experiment in which it was found that particular sequences of speech act types have implications for the structure of the ensuing dialogue and can be correlated with certain aspects of the user, such as his experience in the domain. McKevitt contended that such empirical data, rather than subjective decision-making, should be the basis for constructing user models and argued for the development of automatic techniques for deriving the models.

## Summary

The second workshop was as successful as the first, with all agreeing that subsequent workshops should be held more frequently than at four year intervals. Since the general trend has been for researchers in different areas of user modeling to operate in isolation, such workshops are particularly important as a means of increasing cooperation and cross-fertilization of ideas among the subdisciplines. The Third International Workshop on User Modeling is planned for the summer of 1992 in the German Computer Science Meeting Center at Schloss Dagstuhl, near Saarbrucken, Germany. Program co-chairpersons are Dr. Robin Cohen of the University of Waterloo, Bob Kass of the EDS Center for Machine Intelligence, and Cecile Paris of the Information Sciences Institute. Local arrangements co-chairpersons are Elizabeth Andre, Winfried Graf, and Wolfgang Wahlster, all of the German AI Center at the University of Saarbrücken.

### About the Author

Sandra Carberry is an associate professor of computer science at the University of Delaware. Her research interests include discourse understanding, user modeling, planning and plan recognition, and intelligent natural language interfaces, and she is the author of a new book entitled *Plan Recognition in Natural Language Dialogue* that is part of the ACL-MIT Press Series in Natural Language Processing.

AI Magazine Volume 11 Number 5 (1990) (© AAAI)

# The First International Workshop on Human and Machine Cognition Pensacola, Florida. Topic: The Frame Problem

*Eric Dietrich*

In 1859 Charles Darwin published *The Origin of Species* and exposed the complex and fascinating mechanism underlying speciation. Before Darwin's book, we were in fact ignorant not only of the mechanisms underlying speciation, but of the fact that speciation had occurred and was still occurring. Indeed, most supposed that the species that were on the planet at that time were immutable, and that they had existed from the beginning of the Earth. A profound problem—how to tailor an organism to suit a particular environment—was being solved right under our noses. After Darwin's book, and in part because of it, a new field emerged whose cumulative insights over the years have revealed many of the mechanisms involved in evolution.

In 1877 the Italian astronomer Giovanni Schiaparaelli announced the existence of *canali* on Mars: a network of straight and curved lines running across the planet. *Canali,* meaning channels or grooves in Italian, was translated by the English press into "canals," and with that an intense love affair with Mars and its inhabitants began. Of course, the affair had its share of dark misgivings: what if the inhabitants should turn out to be smarter than we with the same bent for conquering and enslaving? At the center of all this was Percival Lowell, the builder of the Lowell Observatory in Flagstaff, Arizona. He championed the seemingly profound problem of the Martian canals and the nature of the beings on Mars. He spent his life trying to unravel these mysteries. The image of Mars as a planet inhabited by an ancient, canal-building civilization became the popular vision, even though many other astronomers could not find the canals Lowell observed regularly and frequently. Today we know that there are no canals on Mars, and that Lowell and others had seen illusions and atmospheric distortions.

In 1969 John McCarthy and Patrick Hayes formulated the *frame problem:*

> ...in proving that one person could get into [a phone] conversation with another, we were obliged to add the hypothesis that if a person has a telephone *he still has it* after looking up a number in the telephone book. If we had a number of actions to be performed in sequence, we would have quite a number of conditions to write down that certain actions *do not change* the values of certain [propositional functions describing situations] (McCarthy and Hayes, p. 487, emphases added).

More generally, the frame problem is the problem of blocking the vast number of inferences about what has not changed as the result of performing some action *A* while allowing the small number of inferences about what has changed as a result of *A*.

For some of us, the "Frame Problem Workshop" (as it was called) was an opportunity to discuss a methodological question which has become important in AI and cognitive science: Is the frame problem profound or a mistake? Which of the above episodes from the history of science, the "evolution" episode or the "Martian canal" episode, does the frame problem most closely resemble? To some researchers, solving the frame problem will unravel the secrets of the higher cognitive processes and intelligence. To others, the frame problem is much ado about nothing; its very generation rests on seeing something which isn't there. Many of us who are interested in this methodological question have been influenced by Jerry Fodor's book *The Modularity of Mind* (1983). Fodor (who did not attend the workshop) thinks the frame problem is deep and profound.

For still others, the workshop had a different purpose. There is a middle position between the two mentioned above. One can view the frame problem as an interesting, challenging problem arising in certain formalizations of problem solving. Unfortunately, these formalizations follow from a natural, seemingly innocuous, assumption about problem-solving processes: namely that thinking is inferencing. Those who hold this position view the frame problem as a technical problem which together with several other hard, technical problems may require a nonstandard logic for its solution.

All three positions were well represented at the Frame Problem Workshop. For those taking one of the two polar positions ("Martian" or "Darwinian"), the workshop was an opportunity to argue for various philosophical theses regarding the frame problem, its various formulations, and its consequences for AI. For those taking the middle position, philosophical and methodological questions concerning the frame problem were secondary. These researchers were primarily interested in presenting solutions to the frame problem within certain formalizations of problem-solving.

## Some Workshop Details and History

The workshop took place from May 11 to 13, 1989, at Pensacola Beach, Florida. The workshop site incorporated the facilities of both the Dunes Hotel and the neighboring Holiday-Inn. Support was provided by AAAI, the National Science Foundation, Florida High Technology & Industry Council, BDM Corporation, Monsanto Corporation, and The University of West Florida. The workshop chair was Ken Ford of the University of West Florida. The program committee was J. Adams-Webber (Brock University), N. Agnew (York University), F. Brown (Univ. of Kansas), F. Petry (Tulane), L. Reeker (BDM Corp.), and R. Yager (Iona College).

About 40 people were invited to attend. Although many of the attendees hailed from university computer science departments and corporate AI centers, the psychology and philosophy communities were also well represented. Any time the discussion got too confusing or vociferous, the Gulf of Mexico and the white sands of Santa Rosa Island provided opportunities to collect one's thoughts and achieve some measure of repose.

Part of the success of the workshop was due to its format—there were planned discussion groups (for the first afternoon and the third morning) and the parallel sessions were kept to a minimum (occurring only on the second afternoon). All papers were held to a reasonable length allowing plenty of time for questions after the papers.

This was the second workshop on the frame problem. On April 12-15, 1987, the first workshop on the frame problem was held at the University of Kansas (see Brown, 1987). The 1987 workshop focussed on the frame problem in its more narrow original formulation and primarily featured papers on logic-based approaches to solving the frame problem such as circumscription and modal logic.

In addition to providing a forum for the presentation and continued discussion of research resulting from seeds planted at the first workshop, the second workshop consciously expanded its scope to include investigators in cognitive science, philosophy, and other related areas.

## Some of the Papers and Debates

Daniel Dennett, a philosopher from Tufts University, gave the inaugural talk. His paper was entitled "Framing the Question." Dennett is a (mild) "Darwinian." He thinks that the frame problem is important and that figuring out how organisms cope with it is crucial to developing useful knowledge representations for planning and control. He views the frame problem as the problem of predicting the future. All organisms have the frame problem; they must somehow answer the question "Now what do I do?" One answer to this question takes the form of "act randomly and hope for the best" (Dennett says all organisms do this at some time or another). The preferred answer to this question, though, is "represent the world and use your representation for planning." Organisms implement this answer more or less well, ranging from those who represent almost none of the world and are content to let the immediate world and proximal future warn them (e.g., jellyfish) to those who represent a lot of the world and try hard to peer deep into the future,letting the non-immediate future warn them. Dennett argued that (the received doctrine not withstanding) organisms in the latter category do not use logical inference to peer into the future. Instead they oversimplify and self-monitor, learning and using what heuristics they can along the way. As is apparent, even creatures like us, therefore, have to hope for the best.

Pat Hayes then labored mightily to dispel the philosophical spell of profundity that surrounds the frame problem. His paper was entitled "What the Frame Problem Was and Isn't." For the record, Hayes thinks philosophers exercised about the frame problem have made a mish-mash of his and McCarthy's original definition. As Hayes has repeatedly pointed out, strictly speaking, the frame problem arises only in the *situation calculus* (see McCarthy and Hayes, 1969, for a discussion of the situation calculus). In this view, the frame problem is simply the proliferation of frame axioms needed to handle the problem of determining what facts about the world remain unchanged when some action is performed. Somehow, the phrase "the frame problem" came to refer to the more general problem of determining what facts about the world remain unchanged when some action is performed. This is how the term is used today. (This latter point is made in a footnote in Leora Morgenstern's paper. See below.)

However, Hayes does not take the Martian view of the frame problem, but is instead part of the middle group. He summed up the middle position nicely by pointing out that most changes in our environment, either those made by us or some other force, have a moderate, local consequence, but that all our formalisms allow for global consequences that are not moderate. Zenon Pylyshyn used this opportunity to point out that physics has a notion of locality, but that cognitive science and AI do not. He suggested that some smart researcher develop a metric for localizability.

Don Perlis, in "Intentionality and Defaults," related intentionality (the aboutness or meaning of internal symbols), default reasoning, and the frame problem via a single, important (but frequently overlooked) cognitive capacity: the ability to distinguish reality from appearance. He argued that this capacity is a crucial aspect of intelligence. Specifically, an intelligent system needs the ability to process or use two kinds of symbols: those which are primarily conceptual and do not represent anything in the external world (external to the organism or system), and those which do represent something in the external world. But this is not all. The system itself must know which of its symbols are which. That is, the system must know when some symbol represents things as they might have been, used to be, or could never have been, and when some symbol represents things as they are (or at least purports to represents things as they are). Perlis then made three claims: 1) Such a cognitive capacity might explain why we humans have original intentionality rather than merely derived intentionality (if we do); 2) this capacity (and the distinction between reality and appearance) should be influencing research in default logics (Perlis is exploring this line of research), and 3) this capacity and this distinction would help us understand and come to terms with the frame problem.

Ron Loui's paper, "Back to the Scene of the Crime, or Who Survived Yale Shooting?," was the clever, entertaining paper of the workshop. Loui had noticed an interesting connection between the frame problem and the Yale shooting problem (introduced in 1986 by Hanks and McDermott). The Yale shooting problem is a problem in interpreting standard nonmonotonic logics. To quote from Loui's paper "A gun is loaded. We wait. It's fired [at Fred]. Does Fred die because, by default, the gun persisted in being loaded during the wait? Or does Fred persist in being alive, by default, which entails that the gun became unloaded? [The Yale shooting problem] is a prediction problem because when we write the axioms the way Hanks and McDermott wrote them, it is unclear what prediction ought to be made." Loui's position is that a solution to the frame problem must be able to handle the Yale shooting problem as well. He then went through the positions of each of the authors in *The Robot's Dilemma* (Pylyshyn, 1987) and analyzed how their proposed solutions, or other method of dealing with the frame problem, fared when considered in light of the Yale shooting problem. Those authors that did not fare well, Loui declared "dead", victims of the Yale shooting.

Janlert argued that the frame problem is the task of constructing a stable representation of the world. He sees the frame problem as more general than the problem of representing change (the problem of representing change, or relevant changes, is almost universally regarded as the quick and dirty statement of the frame problem). Janlert, therefore, regards the frame problem as a profound problem confronting AI researchers. Janlert separated the frame problem (which he defined as the problem of constructing a representational system that makes the world appear as stable as possible) from several other problems. Three of the these other problems are the prediction problem (how to constrain representations to be true of the world), the revision problem (how to fix one's beliefs when they go awry using a minimum of effort), and the relevance problem (how do we keep search time down to acceptable levels). Janlert did not claim that all of these problems can be solved independently, in fact they probably cannot, he asserted. But he did argue for distinguishing between them on the grounds that doing so will keep the central target in sight: the representation of change. The frame problem and all of its cousins flow from this one central problem, according to Janlert.

The next speaker, Robert Cummins, contended that the frame problem is mitigated when one can exploit principled restrictions on relevant information. This was illustrated by a discussion of Pathfinder, a system that learns conventions governing the use of arbitrary symbols and constructions in communication. Cummins suggested that by embedding a difficult communication/learning problem in a containing coordination problem, the task becomes simple because the embedding provides a clear criterion of relevance, hence puts limits on the information that must be retrieved and maintained.

The afternoon of the first day was spent in one of three discussion groups. Pat Hayes was the group leader of the group I attended (the other group leaders were Zenon Pylyshyn and Clark Glymour). We spent some time trying to figure out what those who say the frame problem is profound could possibly mean, and more generally, why there seems to be such disagreement about the *nature* of the frame problem, rather than about various proposed solutions to the problem, which would seem more natural. Each of us offered a guess, and the others would argue against it. We got nowhere. Just when we would begin to feel certain that the frame problem was much ado about nothing, someone would couch the problem in a way that made it seem profound. We finally gave this up and spent the rest of the time trying to devise some psychological experiments to measure how extensively a subject's knowledge base would change, given certain changes in the subject's environment. Then we broke up for the day.

We began day two with a paper from a member of the middle group—Leora Morgenstern. She discussed the relation between action theory and epistemology. She has discovered two new versions of the frame problem which arise when one's theory of action allows for multiple agents. Morgenstern referred to them (i.e., the two new versions) as the Third Agent Frame Problem and the Vicarious Planning Problem. She has developed a nonmonotonic temporal logic called Motivated Action Theory, which can handle the Third Agent Frame Problem, but she has not yet solved the Vicarious Planning Problem.

Yoav Shoham was next; his paper was entitled "Time for Action." He distinguished two approaches to the study of time and action in AI: 1) change-based systems, in which the concepts of time and action are left implicit as in the situation calculus (McCarthy and Hayes, 1969), and 2) temporal logics that view action as something taking place in time, but which explicitly define a notion of time. In the former view, action is handled intuitively correctly, but such notions as durations and overlapping events are difficult or impossible to represent. In the second view, the notion of action is otiose; it ceases to be even implicitly required. Shoham then proposed a third approach which retains the best advantages of the other two, but without their flaws. This third proposal combines action, time, and knowledge: actions are associated with choices in time which are associated with knowledge (actions are branch points on a time-line which are made on the basis of what is known and affect future knowledge). The key notion is an *evolving time-line structure* which describes an author's epistemic evolution over time.

The next paper, "The Frame Problem, the Modularity Thesis, and Computing," was by Eric Dietrich and Chris Fields. Their paper explicitly confronted Fodor's formulation of the frame problem in his *Modularity of Mind* (1983), and his arguments there for making the frame problem one of the central problems in cognitive science. Dietrich and Fields are "Martians." They believe that casting the frame problem in the role as "most profound problem in the cognitive sciences" is a mistake, a mistake which depends on not taking the hypothesis that thinking is computing seriously enough. They deny that "solving" the frame problem would unravel many of the mysteries in cognitive science and AI. They do, however, think that discussions of it shed light on important issues in the cognitive sciences.

In "The Modal Quantificational Logic Z: A Monotonic Theory of Nonmonotonic Reasoning,"Frank Brown presented Z, a first-order, quantified, monotonic, modal logic. Brown claims, however, that Z can be used to model nonmonotonic reasoning, and Brown advocates using Z for just this purpose. Brown argued that the trick is to represent the validity of sentences as the necessity of the meanings of those sentences. This allows the semantics of a nonmonotonic system to be embedded in Z. Brown explained how this works by showing that several popular alternative systems of nonmonotonic reasoning can be modeled within Z (e.g., he argued that autoepistemic logic, constructive default logic, parallel circumscription and BNO logic can be directly modeled in Z).

Sarit Kraus presented "Nonmonotonicity and the Scope of Reasoning,S a paper co-authored with David Etherington and Donald Perlis. She began with some examples of the failure of circumscription. (The results of this paper are restricted to circumscription, but the authors suggest that their results apply to other formalizations of nonmonotonic reasoning.) The general problem is this: In any particular instance, a default inference is probably reasonable, but in the long run, conditions are bound to change in such a way as to make the inference false. Since we know this, in the long run the inference is not reasonable. So how do we restrict default inferences to the cases where they're reasonable? The frame problem is an example of this: if some condition persists for a certain time interval, it is very often reasonable to infer that the condition persists for a bit longer than the given time interval, even though we know almost nothing persists forever, and hence that the inference is not reasonable all the time. Of course, this is the very problem nonmonotonic logic was designed to solve; when the default inference is erroneous, back up and change the axioms. The authors' point is that default reasoning is over zealous (i.e., it applies the default rules too broadly, thereby making default inferences where they are not, in fact, reasonable), Their solution is to add to circumscription in particular (and to nonmonotonic reasoning, in general) a notion of *scope*. Scope is defined in terms of the *a priori* relevance of a given situation to a particular episode of reasoning. *A priori* relevance is narrow scope and *a priori* irrelevance is wide scope.

In the ensuing discussion, Jack Adams-Webber and Neil Agnew mentioned Ford's (1989) recent paper on Von Wrights's "range of relevance" and Kelly's "range of convenience." They suggested a logical parallel between these concepts and Kraus et al.Us notion of scope, and possibly Pylyshyn's desire for a metric of localizability.

The final speaker of the second morning was Erik Sandewall. His paper was entitled "On The Variety of Minimization Related Conditions for Reasoning About Actions and Plans." Sandewall is interested in unifying temporal reasoning, knowledge-based planning, and qualitative reasoning. He is working in a kind of "blocks world" called "particles world." The simplest world he described contained two "particles" (they were only two-dimensional) moving back and forth from opposite ends on a horizontal line segment, possibly colliding and bouncing back to their starting positions. The line contained a trap door (we are dropping the scare quotes; in particles world, it is a real trap door). Beneath the door was a shaft with rigid sides. The door could open allowing one or both of the particles to fall down the shaft. The object of the game was to make some observations and use them to predict the behavior of the particles and plan actions to accomplish certain goals (e.g., letting particle-1 fall down the shaft but not particle-2). Equipped with this world, Sandewall posed an open problem: how is the set of observation compatible theories related to an appropriate plan for achieving a given goal? This problem is hard because 1) of course, there are many theories explaining the various axioms and laws one might want to use to describe particles world and 2) actions of agents are included in particles world and therefore must be expressed in the theories. and the actions are not determined ahead of time. Observations 1) and 2) simply make it hard to figure out what is going to happen next. In this, particles world is something like our own.

In the afternoon, we had two parallel sessions. There were a total of twelve papers during the afternoon. Not all of them will be reviewed here.

James Fetzer ("The Frame Problem: Scorekeeping with Maximal Specificity") argued that the frame problem can be solved in principle only for closed systems involving finite sets of properties.

Lynn Stein ("An Atemporal Frame Problem") demonstrated that the frame problem is not exclusively temporal. This was interesting because virtually all versions of the frame problem are couched in terms of change and persistence over time. The atemporal frame problem involves counterfactual reasoning: if kangaroos didn't have tails, they would topple over; but if they didn't have tails they might not topple over because they might be good with crutches. That kangaroos might be good with crutches is obviously an assumption that is too far-fetched. Determining that it is too far-fetched is analogous to the frame problem.

Jay Weber ("The Myth of Domain Independent Persistence Reasoning") addressed the problem of proving that a given property persists over time. The strategy is to prove that every action that could change the property did not occur over the relevant time interval. Of course, this strategy is not implemented literally. Rather, the default inference "all actions not known to occur are considered not to have occurred" is used. (There are other versions of this strategy, but they are not theoretically interesting, as Weber pointed out.) Weber then argued that this approach is defective because it assumes a domain-independent mechanism for reasoning about persistence, and such a mechanism is too strong; it is intuitively implausible and produces intuitively wrong answers in versions of the Yale shooting problem.

Josh Tenenberg ("The Robot Designer's Dilemma") made an important, but underutilized distinction. He claimed that the frame problem is properly the designer's problem. It is the designer of a system who has to figure out how to implement a system that can draw relevant inferences about the future state of its world. The system itself doesn't have this problem. How could it? Once implemented, the system executes the specified algorithm. The algorithm determines what the system is going to consider relevant for any particular problem confronting the system. Even if the system changes its mind and considers some other piece of information relevant, this, too, is determined by the algorithm.

Day three started with a continental breakfast after which the groups reconvened to discuss the previous day's plenary papers and continue dialogue started in earlier group meetings. The group leaders for this session were Pat Hayes, Donald Perlis, and Clark Glymour. As may be expected, the discussions were lively and wide ranging, interrupted only by an impromptu dolphin show right outside our window.

After lunch, a plenary panel session convened in which the discussion group leaders were joined by Daniel Dennett and Yoav Shoham to form an interdisciplinary panel. Dennett summed up his impressions of the workshop and argued once again that the frame problem was quite important. And Hayes denied this...again, trying once more to distinguish the frame problem from the general problem of deciding what to do next and what beliefs to update, and arguing that logic was useful to these latter problems. Positions were reiterated. Points were made. Distinctions were drawn. Aspersions were cast (good-naturedly, of course). Confusion reigned. Cummins (in the audience) pointed to his watch. And Glymour suggested halting. Immediately following the plenary panel session Ken Ford made some closing remarks, thanked the participants, and declared the workshop closed.

## Conclusion

I thought the workshop was valuable and productive (and this is not just because I found several allies for the "Martian" view). The discussions of all the different approaches to logic in AI, the high caliber philosophical discussions, the white beaches, and dashing off notes late at night while listening to the waves from the Gulf of Mexico all made for an interesting and enjoyable workshop. Given the enthusiasm of the discussions, and the participation in the evening cruises and boardsailing, this sentiment was shared by all.

Although the workshop was notable for the amount of useful communication between disciplines, it was observed that most of the presentations had either a predominantly computer science (logic) flavor or a philosophical/cognitive science flavor. Thus it was decided, rather than publish an ordinary "proceedings", to prepare two volumes of high-quality refereed chapters arising from the papers and discussions at the workshop. The volume with a computer science orientation will be published by JAI Press and is titled, *Reasoning Agents in a Dynamic World: The Frame Problem,* editors, Kenneth M. Ford and Patrick J. Hayes. Ablex Publishing Corporation is producing the other book (more philosophically oriented) entitled, *The Robot's Dilemma Revisited: The Frame Problem in Artificial Intelligence,* editors, Kenneth M. Ford and Zenon W. Pylyshyn.

The Workshop on Human and Machine Cognition will convene every other year. The topic for spring 1991 is Android Epistemology.

## References


Brown, F. (1987). *Proceedings of the 1987 Workshop on the Frame Problem in Artificial Intelligence.* Morgan-Kaufmann.

Fodor, J. (1983). *The Modularity of Mind.* Cambridge, MA: MIT/Bradford.

Ford, K. M. (1989). RA constructivist view of the frame problem in artificial intelligenceS, Invited Commentary on Agnew and Brown's 'Foundations for a Model of Knowing: (1) Constructing Reality and (2) Fallible but Functional Knowledge,' *Canadian Psychology,* (April).

Hanks, S. and McDermott, D. (1986). Default reasoning, nonmonotonic logics and the frame problem. In *Proceedings of AAAI–1986.*

McCarthy, J. and Hayes, P. J. (1969). Some philosophical problems considered from the standpoint of artificial intelligence. In B. Meltzer and D. Michie (Eds.) *Machine Intelligence* 4. New York: Elsevier, pp. 463-502.


### About the Author


Eric Dietrich is assistant professor of philosophy in the Program for Philosophy, Computers, and System Science at the State University of New York, Binghamton, New York. He is a co-editor (with Chris Fields) of the Journal of Experimental and Theoretical AI. His main research interests are modeling analogical reasoning, and debunking the view held by philosophers such as John Searle and Jerry Fodor that computers are formal symbol manipulators. He thinks that anyone who denies that the brain is a kind of computer is a closet dualist.