REASONING ABOUT SENSING ACTIONS AND ITS APPLICATION TO DIAGNOSTIC PROBLEM SOLVING

by

SON CAO TRAN

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Computer Science Department

THE UNIVERSITY OF TEXAS AT EL PASO

MAY 2000

UMI Number: 9971353

UMI®

UMI Microform 9971353

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

REASONING ABOUT SENSING ACTIONS AND ITS APPLICATION TO DIAGNOSTIC PROBLEM SOLVING

SON CAO TRAN

Computer Science Department

APPROVED:

Chitta Baral, Chair

Michael Gelfond

Vladik Kreinovich

Alessandro Provetti

Mohamed Amine Khamsi

Associate Vice President
for Graduate Studies

THE UNIVERSITY OF TEXAS AT EL PASO

Date: **May 2000**

Author: **SON CAO TRAN**

Title: **Reasoning about sensing actions and its application to diagnostic problem solving**

Department: **Computer Science**

Degree: **Ph.D.** Convocation: **May** Year: **2000**

Permission is herewith granted to The University of Texas at El Paso to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

To my wife Nguyen Thi Hong Thanh,
my daughters Tran Thu Trang and Tran Thanh Tra,
and my parents.

Acknowledgements

I would like to thank Chitta Baral for being my advisor for the years I have been a graduate student in El Paso. Without his many ideas, suggestions, and constant support during this period, many parts of this thesis would not be possible or would be in an unreadable form. I am indebted to Michael Gelfond for being the mentor who can afford losing a lot of time to teach me and other students how to write proofs, how to ask questions, and most of all – how to think scientifically and how to do research. I am grateful to Phan Minh Dung, who encouraged me to enter this endeavor. Without these three men my studies would not have been possible and I would not be where I am today.

I would also like to thank my other committee members Vladik Kreinovich and Mohamed Amine Khamsi for their valuable comments on an earlier draft of this thesis. Thanks to Vladik Kreinovich for the jokes he has told me. Whether it was in his class or in the Knowledge Representation Laboratory, they enlightened my studies a lot.

I am grateful to my wife, my daughters, and my mom for their *patience* and *love*. Without their support and encouragement this work would have had to wait some more years before coming to existence, or perhaps, might never have come to exist at all. Trang and Tra, we will visit Disney World very soon!

I wish to thank Monica Nogueira for her friendship, which helps me going through many difficult moments; Aaron Hardesty and Luis Floriano for using the program I wrote for our robot Diablo; Alessandro Provetti and Frank Fernandez for reading and correcting parts of my thesis; and the other members of the Knowledge Representation

Laboratory who have helped in making my stay in El Paso joyful and productive. Thanks to members of the Artificial Laboratory of Computer Science and Engineering at Arizona State University, especially Dr. Subbarao Kambhampati for allowing me to sit in his class and to use the lab; Binh Minh and Xuan Long for many stimulating discussions on planning; and Chi Tuan for asking me questions for which I could never figure out the answers.

I wish to thank Hung T. Nguyen for his valuable advice and constant support for my professional development. Not to forget are the many enjoyable moments in his wonderful house. It is a pity that I could not have him in my committee. Also, I wish to thank my Vietnamese big brothers, John Hoang Trien and Pho Phan and their families, for treating me and my family as members of their families throughout our years in El Paso. I am really fortunate to known them.

Last but not least, I wish to thank everyone who has helped me in completing this thesis.

El Paso, Texas, USA Son Cao Tran

May, 2000

Abstract

This thesis presents a new approach to reasoning about sensing actions and its application to diagnostic problem solving.

We begin with the definition of an action description language $\mathcal{A}_K$ that allows reasoning about actions and their effects in the presence of incomplete information and sensing actions. To define the semantics of $\mathcal{A}_K$, we introduce a notion of a *combined state*, which plays the same role of *state* in reasoning about actions when complete information about the environment is available. The semantics of $\mathcal{A}_K$ is then defined by transition functions, which map pairs of actions and combined states into combined states. We prove the equivalence between $\mathcal{A}_K$ and other approaches to reasoning about sensing actions such as the situation calculus approach of Scherl and Levesque and the high level action description language approach of Lobo et al. We compute the entailment relation defined by domain descriptions in $\mathcal{A}_K$, denoted by $\models_{\mathcal{A}_K}$, by translating domain descriptions into semantics equivalently extended logic programs. Since the search space associated to $\models_{\mathcal{A}_K}$ is very large, several sound approximations of $\models_{\mathcal{A}_K}$ are proposed. These approximations differ from each other by the number of levels in which reasoning by cases is done.

We argue that actions and narratives play an important role in diagnostic problem solving. In formalizing diagnostic problem solving, we extend $\mathcal{L}$, a high-level action description language for specifying and reasoning about narratives, with static causal laws, sensing actions, and obserable fluents. We also define a notion of diagnostic plans, whose goal is to single out a diagnosis among multiple alternatives.

Table of Contents

Chapter 1

Introduction

1.1 Reasoning about Actions

Representing actions and reasoning about their effects has been an important topic in Artificial Intelligence (AI) since its inception. Research in reasoning about actions is concerned with the development of general formal models of action and the investigation of their usefulness. Studies in reasoning about actions have addressed a broad range of issues:

1. **Knowledge representation:** What formalism can be used to specify actions and their effects ? What are the advantages and/or disadvantages of a particular representation formalism?

2. **Hypothetical reasoning:** What will be true after the execution of a sequence of actions from a particular state?

3. **Planning:** What sequence of actions (or plan) will achieve a goal G from a given state?

4. **Narrative assimilation and explanation:** What had happened that explains the observations about the past? Given a (possibly incomplete) history of a system, what does (and does not) hold in the current state?

In general, an action theory is a set of propositions written in a specialized language that was developed for representing and reasoning about actions and their effects. The semantics of an action theory is then defined by an entailment relation that determines what will be true after a sequence of actions is executed from a given state.

To date, there are several approaches to reasoning about actions. They differ from each other by the representation formalism as well as the techniques used in answering hypothetical queries and planning. Some of the queries to a "action representation" may concern the current state of the domain, the existence of a plan achieving a given goal, or the failure of some components etc.

Approaches to actions can be divided roughly into two groups: one emphasizes the use of mathematical logic in knowledge representation while the other has its basis in operation research. In the first group, we find formalisms that use high-level action description languages [95, 37], first-order logic [80, 84, 103, 105, 116], logic programming [62, 31, 115, 107, 116, 28], or temporal logic [110, 112]. Approaches falling in the second group use Markov Decision Process and its extension, Partial Observable Markov Decision Process (see e.g. [15, 51, 100]).

This thesis follows the mathematical logic approach to formalize actions and their effects, thus emphasizing the logical base of knowledge representation.

The first and perhaps most difficult problem in using mathematical logic to formalize actions, is the *frame problem*. It was first discussed by John McCarthy in [84]. In short, the frame problem is the problem of describing, in a coincise way, the non-effects of actions, i.e., to express what *does not change* after an action is executed. Solutions to the frame problem did not come quickly. In fact, it took about two decades before Pednault presented a solution to the frame problem in situation calculus [95]. Also Baker [1] and Reiter [103] provided solutions to the frame problem in situation calculus. Gelfond and Lifschitz, with the language $\mathcal{A}$ [37], presented a solution to the frame problem using logic programming. It is worth noting that studies on the frame problem have lead to the creation of a sub-field of AI: non-monotonic reasoning and the development of a number of non-monotonic reasoning formalisms such as circumscription [83], autoepistemic logic [92], logic programming [129, 74, 59], and default logic [101].

After solutions to the frame problem were discovered, research in reasoning about actions has moved forward to discuss various aspects of dynamic domains such as the static relation between fluents, the indirect and non-deterministic effects of actions, and the inexecutability conditions etc. Work in this direction has culminated in a number of extensions/modifications of previous approaches or in new formalisms altogether. We review some of the developments related to the *ramification problem* [39] and the *qualification problem* [82] below.

The ramification problem [39] is concerned with the representation of static domain constraints or indirect effects of actions. The qualification problem [82], on the other hand, is concerned with actions that may not be executable in a certain situation. Ginsberg [39] points out the importance of static domain constraints in reasoning about actions and plans since they can lead to a smaller search space. McCarthy [82] discussed the necessity of distinguish between physical and epistemic impossibility of action execution.

The ramification problem has been approached from different directions. Solutions are found using - among other - the principle of *minimal causation* or the principle of *universal causation*. Minimal causation presupposes that changes caused by an action are minimal. On the other hand, universal causation presupposes that change always has a cause in the description a hand.

Lin [70, 71] addressed ramification and qualification with the minimal causation principle. Using this technique, he can capture reasoning about non-deterministic effects of actions as well. Lin [71] also showed how regression can be used to reason about the indeterminate effects of actions. Baral [2] presented another solution to the ramification and the qualification by translating action theories into disjunctive logic programs. Thielschler [123] discussed qualifications and ramifications in the fluent calculus. Kartha and Lifschitz [55] and Giunchiglia et al. [41] extended the language $\mathcal{A}$ to reason about indirect effects of actions and state constraints.

McCain and Turner [78] developed a causal theory for ramifications and qualifications. Later, they developed a new language for causal theories [79], whose the semantics is based on the universal causation. Causal theories are used to express ramifications, qualifications, and other features of action domains. Along this line, Giunchiglia and Lifschitz developed the language $\mathcal{C}$ for representing and reasoning with static domain constraints, indirect effects of actions, and actions with non-deterministic effects. The language $\mathcal{AC}$ of Turner [126] also allows similar features. Lifschitz [67] later showed that the causal logic of McCain and Turner is also applicable to situation calculus. Also the logic PMON [27] for actions and changes has been extended to address qualification and ramification [46]. A comparison between different treatments of ramification can be found in [113, 46].

Besides the aforementioned problems, there are discussions on several other aspects of action specifications such as actions with non-deterministic effects [2, 97, 55, 111], concurrent actions [73, 5, 6], narratives [90, 98, 7, 3, 52], actions with duration [117, 109, 111, 112, 90, 104], natural events [104], and sensing actions [91, 47, 75, 9, 10].

In this thesis, we will follow the high-level language approach of Gelfond and Lifschitz's language $\mathcal{A}$ to representing and reasoning sensing actions (or knowledge producing actions) in the presence of incomplete information. This language is then extended to narratives and static causal laws to address diagnostic problem solving.

1.2 Reasoning about Sensing Actions

In their landmark paper introducing the situation calculus [84], McCarthy and Hayes discussed the importance of sensing actions for knowledge representation and automated reasoning. They also pointed out that sensing actions do not change the world. Rather, they change the knowledge of a reasoner.

As sensing actions change the knowledge of a reasoner, reasoning about sensing actions could be viewed as reasoning about one's knowledge. Moore argued that agent's knowledge plays an important role in planning and acting to achieve a goal, and that the completeness assumption in early planning systems is unrealistic. He presented a modal logic of knowledge, based on Hintikka's logic of knowledge and belief [49, 50], with axioms that can be used to infer conclusions about one's knowledge. Moore defined a possible-world analysis of knowledge [91] based on Kripke's possible-world semantics [63]. The axioms of his logic can be satisfied by imposing restrictions on the accessibility relation between world models. Then this logic of knowledge is embedded into situation calculus to capture actions and their effects. Especially, he presented axioms that describe the effects of sensing/non-sensing actions on the accessibility relation. Morgenstern [93] generalized Moore's single agent framework to multiple agents reasoning about each other's knowledge.

In Moore's formulation, there is no distinction between the agent's knowledge about the world and the agent's belief about the world. Hass [47] later argued for a clear distinction between the two, since assuming that an agent always knows every logical consequence of its knowledge (logical omniscience) might seem too strong. He proposed a syntactic theory of belief and action and used it to describe the connections between belief and planning under the assumption that beliefs are first-order logic consequences of knowledge about the world.

Scherl and Levesque [114] adapted Moore's formulation to situation calculus. They treated the accessibility relation between world models as an epistemic fluent and combined the solution to the frame problem for non-sensing actions of Reiter [103] and Moore's possible-world semantics for sensing actions to have successor state axioms for this epistemic fluent. Scherl and Levesque proved several important results about their formulation such as: knowledge-producing actions do not affect fluents other than the knowledge fluent; and actions that are not knowledge-producing only affect the knowledge fluent as appropriate. They also showed how regression can be applied to knowledge-producing actions. Levesque [65] used this framework to develop language $\mathcal{R}$ for planning in the presence of sensing actions. Lakemeyer and Levesque [64] combined the situation calculus and Levesque's logic $\mathcal{O}L$ to define a new logic, $\mathcal{AO}L$, of acting, sensing, knowing, and only knowing.

In 1997, Lobo et al extended the language $\mathcal{A}$ of Gelfond and Lifschitz [37] to allow sensing actions and actions with non-deterministic effects. They modified the notion of a situation to be a set of states and defined the semantics of the new language by means of transition functions which map pairs of actions and situations into situations. Parallel to this development, we [9, 10] proposed different approximation semantics for sensing actions. In this thesis, we present a new semantics for this extension and formally compare our approach with the approach of Lobo et al as well as that of Scherl and Levesque.

Sensing actions are not only important for planning as well as for describing agents who act in the world. Again, this has been pointed out first by McCarthy and Hayes [84] and later acknowledged by Moore in [91]. Moore even discussed the need for conditional and iterative actions in plans in the presence of sensing actions. In the past, several planners have been developed that can create plans (to some extent) in the presence of incompleteness, and some of them use sensing actions.

Golden and Weld [43, 42] and Etzioni et al [32] developed planners that can reason about *knowledge goals*. They argued that knowledge goals are inherently temporal, and hence, any language that can not express temporal goals such as UWL [32] would not be suitable for planning with knowledge goals. Further, they suggested that knowledge preconditions are not neccessary for a large class of Markovian domains. They proposed a language SADL [43] for use in their planners and introduced a minimal but extremely useful set of knowledge-temporal goal such as '**satisfy**', '**hands-off**' and '**initially**'. Besides knowledge goals, SADL also allows universally quantified goals. SADL's semantics is similar to the approximation semantics of [9] in that it uses a three-valued logic and does not allow reasoning by cases.

Goldman and Boddy [44] extended STRIPS [33] to planning under incomplete information. To do so, they used a single model of the world to represent the planner's state of the knowledge, in which fluents are either true, false, or unknown. Later, Goldman and Boddy [45] developed a new language for conditional planning, WCPL, that allows context-dependent, non-deterministic, sensing, or even knowledge-destroying actions and distinguished between the planner's information state and the world state. They use propositional dynamic logic to express conditional plans, and reason about information-gathering (sensing) and the agent's information state.

It could be noted that the formal study of sensing actions and the development of planners with sensing actions are often done separately. While most of the available planners for complete information are sound and complete, soundness and completeness of planners with sensing actions are often left unanswered [43, 44, 45].

In this thesis, we present several three-valued semantics for action domains with sensing actions. Moreover, we prove that these new semantics are sound with respect to Scherl and Levesque's semantics, Lobo et al's semantics, and our semantics based on combined states for action domains with sensing actions. Thus, our three-valued semantics may be viewed as sound approximations of the main semantics of conditional planners. One advantage of these approximations is that their state-space is much smaller than that of possible worlds semantics for sensing actions. Therefore, they can be used to build provenly correct, yet efficient, conditional planners.

1.3 Action Theories and Diagnostic Problem Solving

The design of diagnostic systems is a field of Knowledge Representation related to actions, as of recently. There are now two main approaches to designing a diagnostic reasoning systems: the experimental and the "diagnosis from first principles" approach. Results of the first approach are expert systems, where the heuristic information - e.g. rules of thumb, statistical information, or experts' knowledge - is considered.

Diagnosis from first principles (or model-based diagnosis) is based on the idea that knowledge about the internal structure of a system can be used to diagnose about its failures [21, 23, 38, 22]. Reiter [102] provided the first formal characterization of diagnosis, which has been largely adopted for model-based diagnosis.

Diagnosis from first principles assumes that diagnosis of a system starts with a description of that system itself and a set of observations about the system behavior. If the observations are conflicting with the system description, i.e., the system does not behave as it is supposed to, then a diagnosis is needed. A diagnosis is then a set of components which, when assumed to be functioning abnormally, will explain the observations. It is often reasonable to look for the minimal set of faulty components. As such, candidate diagnoses following Reiter's definition are often called consistency-based diagnoses. Abductive explanation, an alternative in characterizing candidate diagnoses, was proposed by Poole [99], Brusoni et al. [16], and Console et al. [19].

A limitation of model based diagnosis lies in the assumption that the systems in consideration are static, i.e., observations are about a single state. Attempts to overcome this limitation have been proposed by Thielscher [124] and McIlraith in some of her works [86, 87, 88]. McIlraith argued that besides diagnosis, diagnostic problem solving must also address testing and repairing. In [87], McIlraith provides a formal account of testing for static domains.

In an earlier note [12], we developed this aspect by noting that diagnosis of a dynamic system must also take into consideration the possible observations along a system history. In this thesis, we will provide a formal characterization of diagnostic problem solving for dynamic domains using an extension of the action language $\mathcal{L}$ of Baral et al [7] with sensing actions and static causal laws.

1.4 Specific Research Contributions

The research contributions of this thesis can be divided into two major parts. The first part consists of an extension of the high-level action description language $\mathcal{A}_K$ with sensing actions and several sound approximations of its entailment relation. It also contains a sound and complete translation of $\mathcal{A}_K$ domain descriptions into disjunctive logic programs. The second part consists of an extension of the language $\mathcal{L}$ with sensing actions, static causal laws, and narratives and a formal account for diagnostic problem solving.

1.4.1 A Transition Function Based Approach to Reasoning With Sensing Actions and Incomplete Information

The high-level action description language $\mathcal{A}$ was introduced by Gelfond and Lifschitz [37] for representing and reasoning about actions. Its distinguished feature is an independent semantics, based on transition functions which map pairs of actions and states into states. Since $\mathcal{A}$ is developed with the intention of having a minimal language for reasoning about actions, it does not allow other important features of action specifications such as static state constraints, actions with non-deterministic effects, concurrent actions, sensing actions, etc. Other researchers have extended $\mathcal{A}$ or developed variations of $\mathcal{A}$ to allow such features. We have discussed this in Section 1.1.

As in all extensions of $\mathcal{A}$ that capture new features of action specification, extending

$\mathcal{A}$ to allow sensing actions requires a syntactic definition of new propositions for representing sensing actions and an extension of the transition function to incorporate the additional features. The syntactic extension to allow sensing actions is done in the straightforward manner: we add to $\mathcal{A}$ a new proposition type to describe sensing actions and their effects[1].

Since the semantics of action description languages is defined based on the transition functions, to extend $\mathcal{A}$ to allow sensing actions, the first question we investigate is: "*What is a state in the presence of incomplete information?*"

High-level action description language with sensing actions has been discussed previously only by Lobo et al [75]. In their approach, they define the notion of a *situation*, which is a set of world states representing the possible world models an agent think he might be in. A transition function is then defined as a mapping from pairs of actions and situations into situations. One disadvantage of this approach is that domain descriptions with sensing actions might have many transition functions, even when they contain only deterministic actions. In some cases, this is counter-intuitive. We will discuss this issue in details in Chapter 7.

In this thesis, we answer the question above by defining the notion of a *combined state*, which is a pair made up of a world state and a set of states representing the possible world states the agent thinks it might be in. In fact, the second component is a situation in the notation of Lobo et al.. We then define the transition functions of domain descriptions as mappings from pairs of actions and combined states into combined states. With this definition, we avoid the problem of having counter intuitive transition functions as in [75].

We also explore another view of a state in the presence of incomplete information, namely, as a three-valued state. We define an *approximation state* as a pair of disjoint sets of fluents $\langle T, F\rangle$ to represent the agent's knowledge about the world. Intuitively,

[1]The same is done by Lobo et al [75].