科技资料

ARTIFICIAL INTELLIGENCE AND COMPUTER VISION

ARTIFICIAL INTELLIGENCE AND COMPUTER VISION

Proceedings of the Seventh Israeli Conference
26-27 December, 1990, Ramat Gan, Israel

Edited by

Yishai A. FELDMAN

Tel Aviv University
Israel

Alfred BRUCKSTEIN

Technion - Israel Institute of Technology
Israel

1991

NORTH-HOLLAND

AMSTERDAM • OXFORD • NEW YORK • TOKYO

ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
655 Avenue of the Americas
New York, N.Y. 10010, U.S.A.

Library of Congress Cataloging-in-Publication Data

Artificial intelligence and computer vision / edited by Yishai A. Feldman, Alfred Bruckstein.
p. cm.
Papers from the Seventh Israeli Symposium on Artificial Intelligence and Computer Vision, held Dec. 26-27, 1990, Ramat-Gan Israel.
Includes bibliographical references.
ISBN 0-444-89126-9
1. Artificial intelligence--Congresses. 2. Computer vision--Congresses. I. Feldman, Yishai A., 1958- . II. Bruckstein, Alfred Marcel. III. Israeli Symposium on Artificial Intelligence and Computer Vision (7th : 1990 : Ramat-Gan, Israel)
Q334.A7625 1991
006.3--dc20 91-3265
CIP

ISBN: 0 444 89126 9

This book is printed on acid-free paper.

Printed in The Netherlands

FOREWORD

The Israeli Symposium on Artificial Intelligence and Computer Vision continues an already established tradition. This is the seventh such conference, held in Tel Aviv, Israel. Researchers from many places in the world and their Israeli colleagues gathered in December to share some of their recent results, and to discuss current reserarch in Artificial Intelligence and Computer Vision.

We are happy to note that a great many papers presented in this volume are from researchers abroad, showing the growing interest of the international community in the work done in Israel. A further mark of this interest, as well as fuel to grow it, is the fact that this year, for the first time, the proceedings of this conference will be distributed internationally by Elsevier Science Publishers, Amsterdam.

AI and Vision are very popular with computer science and engineering students. Industry is also very interested in finding out the possibilities for lucrative applications. We are always happy to share our results with them. We hope all the participants will enjoy this event, and will learn some things from it.

The papers presented here were selected by the organizing committee from a large number of submissions, through a standard refereeing procedure. The accepted papers are very interesting and the subjects covered are varied, reflecting most current trends of research in the field!

Yishai A. Feldman
Tei Aviv University

Alfred Bruckstein
The Technion

PROGRAM COMMITTEE

Dr. Yishai Feldman (Chairman, AI Track) — Tel Aviv University
Prof. Alfred Bruckstein (Chairman, Vision Track) — Technion

AI:

Mira Balaban — Ben Gurion University
Moshe Ben Bassat — Tel Aviv University
Rina Dechter — Technion
Tamar Flash — The Weizmann Institute
Ehud Gudes — Ben Gurion University
Daniel Lehmann — Hebrew University
Marc Luria — Technion
Yoram Moses — The Weizmann Institute
Uzzi Ornan — Technion
Jeff Rosenschein — Hebrew University
Uri Schild — Bar Ilan University
Ehud Shapiro — The Weizmann Institute

Vision:

Zvi Meiri — IBM
Amnon Meizles — Ben Gurion University
Shmuel Peleg — Hebrew University
Moshe Porat — Technion
Micha Sharir — Tel Aviv University
Shimon Ullman — The Weizmann Institute
Michael Werman — Hebrew University
Haim Wolfson — Tel Aviv University
Yehezkel Yeshurun — Tel Aviv University

CONTENTS

Artificial Intelligence and Computer Vision
Y.A. Feldman and A. Bruckstein (Editors)

Structuring Events for their Pronominal Reference

Ethel Schuster

Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104-6389, USA[1]

Abstract

When we as speakers interact with each other using language, we tend to use pronouns to refer to the items that are being talked about. In general, pronouns serve as links between a given item introduced in the discourse and subsequent references to that particular item. In this paper we focus on pronominal reference to events and actions. While one pronoun can refer to a *single* event or action, it may also refer to a *compound* one comprised of several events or actions. We propose an approach where one pronoun can refer to a set of events as long as an underlying intra-structure relationship can be identified among those events. We discuss three different types of intra-structure relationships among events, and the advantages that they provide to pronominal reference.

1 Introduction

When we as speakers interact with each other using language, we tend to use pronouns to refer to the items that are being talked about. In general, pronouns serve as links between a given item introduced in the discourse and subsequent references to that particular item. These items can be of many sorts: objects, people, events, facts, states, or actions. Rather than having to describe a particular object or event every time we want to talk about it, we can make use of a particular pronoun to refer to it. While one pronoun can refer to a *single* event, it may also refer to a *compound* one (comprised of more than one event or action).[2]

We divide the problem of pronominal reference to events into two parts: (1) the intra-structure relationship, which describes how the events are organized into *compound* events, and (2) the relationship between the pronoun and the events themselves (be they

[1] Author's mailing address: 3-5 Apple Ridge Rd., Maynard, MA 01754, USA.

[2] Previous work on reference [15, 7] has focussed primarily on individuals and objects. Lately, the work on reference has been extended to events and actions, [10, 12, 11, 13], as well as propositions [3, 4], and discourse segments [16, 17]. In this paper we focus mainly on pronominal reference to events and actions. We consider actions as being specific instances of situations in which an agent performs the action. Events on the other hand, do not require that the agent be explicitly identified. Throughout the paper, we do not distinguish between events and actions, using events to indicate both unless it is otherwise indicated.

single or *compound*). Events related to one another form different types of compounds. Our approach presents a solution where a pronoun can refer to a set of events as long as an underlying intra-structure relationship can be identified among those events to form a compound one. Our goal is to use this approach in generating text that handles reference to events.

This paper discusses three types of intra-structure relationships among events: sequence, generation, and causation. It describes the advantages that these structures can provide when generating text with reference to events. In section 2 we describe our data. In section 3 we describe and explain the intra-structure relationships. Section 4 presents a formal definition of the intra-structure relationships. Section 5 describes GRECO, a small system that generates text. In section 6 we present a summary.

2 Naturally Occurring Data

The approach presented here is based on the analysis of naturally occurring data. We used two different sets of data, referred to as *the EMACS transcripts* and *the Questionnaires*, as the basis for the work presented here.

2.1 The EMACS transcripts

This set consists of several dialogues between two people interacting with each other, one acting as an expert and the other as a novice [8]. The novice was learning how to use the EMACS text editor by interacting with the expert in English, via a computer terminal. From the point of view of building systems that use natural language for both understanding and generation, the main purpose of collecting these dialogues was to characterize the language used by both the expert and the novice.

These transcripts are considered "task dialogues": both the expert and the novice were working together on a task. In this project, the novice's task was to edit a file on-line so that it would look exactly like a hard-copy of that file which s/he had been given. Ultimately, the goal was to learn how to use EMACS. The expert's task was to help the user as much as possible in editting the file in the easiest way, i.e., with the smallest possible number of commands or keystrokes. Because the expert and the user had a particular task to achieve, they were not particularly concerned with the language they used. This indicates that the language used was spontaneous and it gives us a sense of how different people use language in different ways. For the specific goal of this research, the main concern was to characterize the use and potential interpretation of pronouns referring to events and actions in the dialogues. Additional data to support these results, and from which examples are taken, include the manual for the AMIGA 2000 [2].

2.2 The Questionnaires

The second set of data consists of a number of questionnaires that were given to undergraduate students in introductory classes in Computer Science and Linguistics at the

University of Pennsylvania. Each student was given the following two paragraphs:

(1) Mary thought of getting a PhD in English. *It/That* would have pleased her father and *it/that* would have infuriated her husband. But *it/that* was just what she wanted.

(2) John thought about becoming a street person. *It/That* would hurt his mother and *it/that* would make his father furious. *It/That* was not a clever thing to do.

with either one of the pronouns *it* and *that*. That is, some students got the text with only *it*, some with both *that* and *it*, and some with only *that*. No student got the same pattern of pronouns in both paragraphs. Their task was to answer the following two questions for each paragraph:

(3) (i) What would have infuriated Mary's husband?
(ii) What did Mary want?

(4) (i) What would make John's father furious?
(ii) What was not a clever thing to do?

The initial motivation for running this set of experiments was to test the use and interpretation of the two pronouns *it* and *that* in order to determine how the choice of those two pronouns affected the choice of their referents. Indeed, as the results showed, the different pronouns affected the results in a consistent way. The ultimate goal was to provide a basis for developing an approach to the use of these two pronouns.

3 Intra-Structure Relationships Among Events

We assume that every pronoun in the text refers to single event or to a compound set of events. A pronoun can refer to one single event.[3] This single event can be one of several types including: a specific event as described by the text with a given agent and possibly occurring at a specific time as in example (5) below

(5) John was shot in Philadelphia at midnight. *That* really shocked the people in the city.
That = John being shot in Philadelphia at midnight.

or a more general event obtained by generalizing on the specific one as in example (6)

(6) John was shot in Philadelphia at midnight. *That* would never happen in Boston.
That = someone/John being shot at midnight

[3]This is probably close to the notion of antecedent as characterized by linguists, where the pronoun (anaphoric expression) refers to the exact text as it appears preceding the pronoun, and it can be interpreted by either deleting that text or copying it over to where the anaphor occurs.

Reference to single events such as those in examples (5) and (6) has been discussed in [10, 3, 11].

When a pronoun refers to more than one event, these events form a *compound* event with an underlying internal structure among them. This structure allows for the pronoun to obtain its referent: one pronoun refers to one compound event where the compound is a set of events related to each other by an intra-structure relationship.

3.1 Sequence

A set of events can be related as a sequence if/when the events in the set are described or are performed (by one or several agents) in a linear temporal order. When describing a sequence of actions performed by an agent, we can view the agent doing action B as being subsequent to that agent doing action A if and only if it is correct to say that the agent did A "and then" (or "and later") did B [5], (p. 21). While one event follows the other, it is important to note that they may overlap.

A pronoun can refer to the entire sequence or, as it is shown later on, to parts of that sequence. Consider the following text, where a pronoun can refer to the sequence as a whole:

(7) U: How can I drag an icon?
E: You point to the icon, hold down the Selection button, and move the mouse. A copy of the icon moves and is repositioned at the point where you release the Selection button. Go ahead, try *it*.

In (7) the pronoun *it* refers to a sequence of actions that the user can perform to drag an icon: (1) pointing to the icon, (2) holding down the Selection button (3) moving the mouse, and (4) releasing the mouse. We say that sequence acts as the intra-structure relationship among these events by connecting one to the other. Note that two of the actions in the sequence (moving the mouse and holding down the Selection button) overlap.

We could also view this sequence as a linked list and provide a label for it.

drag icon sequence =
point to icon
↓
hold down Selection button
↓
move mouse
↓
release mouse

Then, when trying to specify what the pronoun refers to, there are at least two descriptions. One could describe this sequence as "dragging an icon" which can be either the label of the sequence (or the name), or the entire sequence of actions that form the list. It is important to stress that this sequence of actions becomes a compound action because of the intra-structure relationship among them.

3.1.1 Individual Events within a Sequence

An individual event within a sequence can also be selected for reference. The choices could be: the first event, the last, or any of the events in between. However, we have found in our analysis of both the EMACS transcripts and the questionnaires that, usually, the one event chosen from a sequence, is the last one unless another one is explicitly specified in the text. In the questionnaires, subjects were given two short paragraphs were the pronouns differed. When given the paragraph

(8) John thought about becoming a street person. *It* would hurt his mother and *that* would make his father furious. *That* was not a clever thing to do.

and the question "What was not a clever thing to do?" the answers were as follows:

43.47%	"making his/John's father furious"
34.78%	"John (thinking about) becoming a street person"
4.34%	"hurting his/John's mother and making his/John's father furious"

When given the paragraph

(9) Mary thought of getting a PhD in English. *It* would have pleased her father and *that* would have infuriated her husband. But *that* was just what she wanted.

the answers to "What did Mary want?" were as follows:

61.90%	"infuriating her/Mary's husband"
4.76%	"Mary (thinking about) getting a PhD"
23.80%	"Pleasing her/Mary's father and infuriating her/Mary's husband"

These results indicate that while it is possible to refer to a sequence of events as described here, it is also possible to refer to the last event in that sequence. However, it is important to mention that the choice of the referent can vary depending on the choice of the pronoun. For instance, when the paragraph in example (8) was given as

(10) John thought about becoming a street person. *It* would hurt his mother and *it* would make his father furious. *That* was not a clever thing to do.

only 16% of the subjects chose "making his/John's father furious" as their answer to the question "what was not a clever thing to do" while 44% chose "hurting his/John's mother and making his/John's father furious."

3.2 Generation

Generation was introduced by Goldman [5] as a way to explain how two actions are related to each other. Generation is a relationship holding between ordered pairs of actions performed by the same agent. It can be described as the relationship in which an agent does B "by" doing A, where A and B are both actions that occur during the same interval of time [5], (p. 20). It is important to stress that in the generation relationship, both actions A and B are specific instances that are performed by the same agent S (in Goldman's terminology, they are act-tokens). In particular, the generation relationship occurs between act-token A and act-token B. If an agent performed one action and

thereby, without any effort on his/her part, did another, then we can say that his/her performance of the former action **generated** the performance of the latter.

For instance, an agent flipping the switch (causally) generates that agent's turning on the light. Another example,

(11) U: I want to select the icon for the Workbench disk.
E: You can select the icon for the Workbench disk by double clicking on it. Try *it*.

The referent of the pronoun *it* can be (1) double clicking on the icon (for the Workbench disk), or (2) selecting the icon (for the Workbench disk), or both (3) select the icon (for the Workbench disk) by double clicking on it. Because the potential referents are described as being in a generation structure with respect to one another, it is possible to consider any of the three as possible referents. The generation structure specifies that "U double clicking on the Workbench icon at a given time can **generate** U selecting the Workbench icon at that given time." This structure "connects" double clicking on the icon and selecting the icon. Having this structure establishes the connection between two potential referents hence preventing the need to commit the referent to being either one or the other. This is an advantage for generating text with pronouns referring to events that can be seen as being in a generation relationship. The compoundd structure allows for any of the potential referents to be chosen as "the referent" while still indicating what the relationship of that referent is to the other potential referent that was not chosen. If the generation relationship between two potential referents is not made clear, they may be seen as independent from one another. Consider

(12) E: Select the icon of the Workbench disk. Double click on that icon and once you have done *that*, the icon will be open.

Here, the user, when reading the instructions as given by the expert, may fail to understand that the expert is referring to "double clicking on the icon for the Workbench disk" as the way to select it. If, on the other hand, the expert provides his instructions as in (13) making the generation structure explicit, it guarantees that the referents can be understood as being related to one another, and therefore be interpreted appropriately.

(13) E: Select the icon for the Workbench disk. You do *that* by double clicking on the icon and then ...

The claim here is that having a generation relationship between two events to form a compound provides some flexibility in choosing the referent of a pronoun. If a pronoun has two potential referents and they are viewed as a generation-related compound, then the referent of the pronoun can be understood as being either one of the events in the compound or the two of them *together*.

3.3 Causation

Causation is defined as an intra-structure relationship between an action and an event, where the action causes an event to happen. When one act causes an event to happen, we

can talk about causation between them. To use Goldman's example, S's act of flipping the switch causes the event of the light's going on. Consider (14) where the pronoun *that* can refer to "being thrown out of the editor," which in turn was caused by U's typing *A*-n. We can then say that the referent in this example can be "being thrown out of editor mode because of U typing *A*-n."

(14) U: I typed *A*-n and I was thrown out of the editor. I did not expect *that* to happen.

It is important to distinguish between generation and causation. In generation, the agent that performs both actions is the same and the actions occur at the same time. In causation, the agent's performing one action causes an event to happen. This event need not occur at the same time as the action performed by the agent. This distinction is expressed in how the referents are specified. In generation, the agent has to be explicitly identified in the referent whereas in causation, if the referent is pointing to the event caused by the action, the agent is not mentioned.

In summary, the representation of events and actions has to allow for the following in order to refer to them pronominally: (1) one single action or event, and (2) a compound event made up of several events related to each other in intra-structure relationships such as: (a) a sequence of actions or events, (b) two actions related by generation, or (c) an action and an event related by causation.

4 Formalizing the Structures

In this section we develop a formal representation of the intra-structure relationships as described in the previous section. Formally, these structures define relationships within events, which permit a pronoun to refer to them.

To represent the intra-structure relationships, we use an interval-based formalism developed by Allen [1]. This is a typed first-order predicate calculus where events and actions are viewed as occurring over intervals of time. The terms in the language are of different types:

- TIME-INTERVAL, which denotes time intervals;
- PROPERTY, which denotes propositions that may or may not hold during a particular time;
- terms corresponding to objects in the domain.

Allen introduces two predicates: HOLDS and OCCUR. HOLDS(p,t) asserts that a property p holds (i.e., is true) during the time interval t. One consequence of this predicate is that a property p holds over all subintervals of t. For example, if "John owned a house throughout 1980", then it is true that he owned a house in January of 1980, as well as February, etc.

OCCUR(e,t) is true if and only if the event e happened over the time interval t, and there is no subinterval of t over which the event happened. Consider for instance the

event in which John ran program *prime*, which occurs over interval t. It is not possible to say that John ran program *prime* over any subinterval of t. We can define a predicate for running a program with two arguments: the agent and the object.

(15) ran(john,program_*prime*).

To assert that John actually ran program *prime*, we need to assert that this event occurred over some time interval $t1$ as in

(16) OCCUR(ran(john,program_*prime*),$t1$).

In addition to these predicates, Allen introduces 13 primitive relations that can hold between temporal intervals [1], (pp. 128-129). Those are the following relations and their inverses:

- DURING($t1$,$t2$): time interval $t1$ is fully contained within $t2$.
- STARTS($t1$,$t2$): time interval $t1$ shares the same beginning as $t2$, but ends before $t2$ ends.
- FINISHES($t1$,$t2$): time interval $t1$ shares the same end as $t2$, but begins after $t2$ begins.
- BEFORE($t1$,$t2$): time interval $t1$ is before time interval $t2$, and they do not overlap in any way.
- OVERLAP($t1$,$t2$): interval $t1$ starts before $t2$, and they overlap.
- MEETS($t1$,$t2$): interval $t1$ is before interval $t2$, but there is no interval between them, i.e., $t1$ ends where $t2$ starts.
- EQUAL($t1$,$t2$): $t1$ and $t2$ are the same interval.

Allen also defines a predicate IN describing the relationship in which one interval is fully contained in another:

$$\begin{aligned} in(t_1, t_2) \quad &\Leftrightarrow \quad (during(t_1, t_2) \\ &\vee \quad starts(t_1, t_2) \vee finishes(t_1, t_2)) \end{aligned}$$

Given this logic it is then possible to define each one of the intra-structure relationships that comprise the compound events.