DISTINGUISHING FACT FROM OPINION AND EVENTS FROM META-EVENTS

Christine A. Montgomery
Operating Systems Division
Logicon, Inc.
Woodland Hills, California, U.S.A.

ABSTRACT

A major problem in automatically analyzing the text of military messages in order to synthesize data base elements is separating fact from opinion, i.e., identifying factual statements versus evaluative commentary such as degree of belief or confidence in the parameters of a particular event, comparative evaluations of entities and events, predictions of future events, etc. This paper describes a model which accounts for the range of factual to evaluative information in the message traffic, and discusses a means for representing such information in the context of an experimental system for automated data base generation.

1.0 BACKGROUND

For several years, with some interruptions, a research and development effort has been in process in the Operating Systems Division of Logicon (and its predecessor company, Operating Systems, Inc.) aimed at the automated creation of data base elements from the text of electronically transmitted military messages.[1] [Kuhns and Montgomery 1973, Kuhns 1974, Kuhns et al 1975, Silva and Montgomery 1977, Silva et al 1979a, 1979b, Dwiggins and Silva 1981.] The objective of this research has been to provide an automated capability to supplement the presently largely manual, labor-intensive task of maintaining the currency of data bases which derive their information elements from the text of military messages. Although some effort has been devoted to primarily interactive approaches to the problem, and to messages which have highly predictable columnar summary formats, the majority of the research and development work has concentrated on the more difficult task of analyzing unformatted narrative text with user interaction limited to occasional assistance to the automated system.

A testbed system called MATRES has been constructed in Prolog to run under the UNIX operating system on the PDP 11/70. MATRES is a knowledge based system for understanding the natural language text of event-oriented messages in the domains of air activities and space/missile (S&M) activities. The knowledge structures in MATRES, called "templates", are essentially frames or scripts describing entities and events, which answer the military user's basic questions about these phenomena, as illustrated in the simplified view of an event template presented in Figure 1.

The templates are hierarchically organized; lower level templates deal with objects or times, mid level with events containing objects and times, higher level with activities composed of events. The slots in the templates contain procedures which operate upon the output of the Definite Clause Grammar (DCG) to instantiate the templates.

We are currently using a corpus of approximately 125 messages in the S&M domain as a basis for developing a scenario for evaluation of the extended MATRES testbed, as well as a testbed for a related knowledge based system, the Active/Introspective Information System [described in Montgomery and Ruspini 1981, and Ruspini 1982] for which MATRES serves as a front end. The scenario involves two simulated nations, the Delta Confederation of the Atlantic States and the Epsilon Republic. Both nations have space programs, and each is interested in monitoring the technological progress of the other, using their own satellite and sensor resources and those of other friendly nations. The set of messages to be analyzed by MATRES are mainly reports of space and satellite launches and orbital activities of the Delta Confederation, which are being monitored and evaluated by the Epsilon Republic. The text of messages used in the scenario has the structure and format of actual messages reporting on S&M activities, although the lexicon is substantially different.

As discussed in several previous technical reports prepared under earlier contracts with RADC ([Kuhns and Montgomery 1973], [Silva et al 1979a], [Silva et al 1979b]), the subset of the English language on which the text of intelligence messages is based is essentially a specialized language for reporting events. Intermixed with factual statements reporting on entities and events, however, is much evaluative commentary. Moreover, press announcements of the Delta Confederation are included in the reports, and evaluative comments are made both about the events reported in the press announcements and the announcements themselves. In synthesizing data base elements from these messages, it is crucial to sort out these different levels of information.

This paper defines an approach to identifying and labeling these types of information so that they can be exploited in the context of MATRES and the data base systems which it serves.

1. This work has been carried out under the sponsorship of the Rome Air Development Center (RADC), U. S. Air Force Systems Command, Griffiss Air Force Base, New York.

INFORMATION PARAMETERS OF A PROTOTYPE TEMPLATE

QUESTION	PROTOTYPE EVENT	PROTOTYPE OF SPECIFIC EVENT	EXAMPLE
WHAT	EVENT TYPE	AIRSPACE VIOLATION	AIRSPACE VIOLATION
WHO	AGENT (OR OBJECT PLUS OWNER)	AIRCRAFT OWNED BY COUNTRY C	A UGANDAN FIGHTER AIRCRAFT
WHEN	TIME OF THE EVENT	TIME AT WHICH VIOLATION OCCURRED	AT ABOUT 0200 HOURS ON 25 APRIL 1978
WHERE	LOCATION AT WHICH EVENT TOOK PLACE	LOCATION AT WHICH VIOLATION OCCURRED	6 MILES FROM THE KENYA BORDER NEAR SUAM
TO WHOM	PATIENT, OR ENTITY AFFECTED BY THE EVENT	OWNER OF VIOLATED AIRSPACE	KENYA
WHY	INTERPRETATION OF THE EVENT	PROBABLE REASON FOR VIOLATION EVENT	PROBABLE RECONNAISSANCE MISSION

Figure 1. Four Aspects of an Event Template.

2.0 THE RESEARCH PROBLEM

As noted above, the sublanguage of the message corpora for the air activities and S&M domains is essentially a vehicle for reporting events. Kuhns and Montgomery [1973] presented a detailed methodology for classifying the various types of events described in the messages, which is shown in Figure 2, and summarized below. Before describing the event classification, however, it is enlightening to review briefly some example messages in order to understand the motivation for this rather complex model.

Some messages -- for example, those encountered in our previous research on the air activities domain -- may report only primitive events. However, as noted above and illustrated in Figure 3, a message may in fact be a report of a report-- that is, it may include a report of an event by some other source than the originator of the message. The "announcement" is thus a report of a "launch" event, which is the basic or primitive event being reported. The "announcement" is an event, but it is clearly not on the same level as the primitive event. Rather, it is a report about the launch, a meta-event that incidentally introduces a new information source of different credibility than the originator of the message.

However, this distinction alone is not sufficient to account for the difference between the initial two sentences of the example message and the third sentence, which contains an evaluation of the announcement, stating that it was characterized by "routine" wording. It is thus an evaluative commentary on the press announcement of the launch event. Since the announcement has been defined as a meta-event, the comment represents another meta-level. In fact, in reviewing additional examples of the message traffic in this scenario, it is clear that, in order to accurately distill and represent information contained in the text of these messages, the analytical methodology must identify and uniquely label the following types of information:

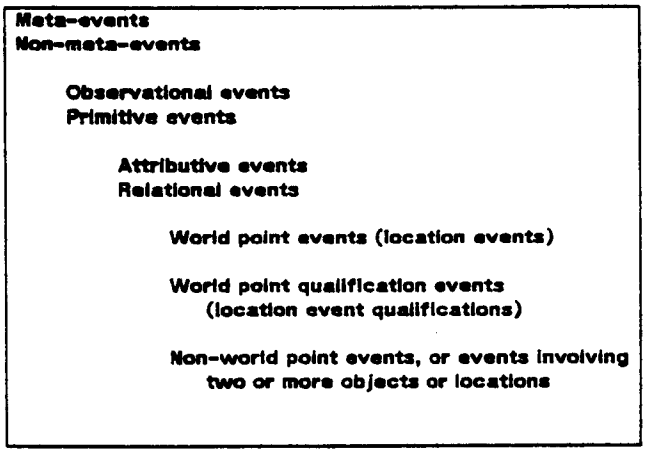

Figure 2. Classification of Events for the Message Sublanguage

NYT ANNOUNCES LAUNCH OF TERREX 558 AND TERREX 559.

AT 1657Z, NYT-THE DELTA NEWS AGENCY-ANNOUNCED THE LAUNCH OF TERREX 558 AND TERREX 559, THE MAN RELATED SPACECRAFT LAUNCHED FROM THE BOGOTA MISSILE TEST RANGE AT 0800Z ON 17 NOVEMBER. THE ANNOUNCEMENT CONTAINED ROUTINE WORDING AND STATED THE ESV'S WERE PLACED IN ORBIT BY A SINGLE CARRIER ROCKET. ORBITAL PARAMETERS CONTAINED IN THE ANNOUNCEMENT ARE:

APOGEE _____ KILOMETERS
PERIGEE _____ KILOMETERS
INCLINATION _____ DEGREES

Figure 3. Example S&M Message.

- factual statements about events made by Epsilon Republic reporters,
- degrees of belief in those statements,
- evaluations of events and predictions of future events,
- degrees of belief in those statements,
- factual statements about events made by the Delta Press,
- degrees of belief of the Epsilon Republic reporters in such statements,
- predictions of future events in the press announcement,
- evaluations of events reported and predictions made by the Epsilon Republic reporters.

3.0 ANALYTICAL METHODOLOGY

This analytical methodology is based on a model of the real world situation which underlies the scenario. Part of the model derives from the event classification schema mentioned above, and illustrated in Figure 2, which was developed to account for the levels of content occuring in the event-oriented message discourse. In this classification, there are two major types of events, meta events and non-meta events. Of the latter, events may be observational, or primitive. An observational event is a direct perception of an event, which may be a visual perception (e.g., "observe", "sight"), or in the case of a sensor, an electronic measurement of the emitted energy characterizing the event. A primitive event is thus a physical event of some kind which does not involve an observation or perception. Primitive events may be attributive or relational. An attributive event describes a situation in which a particular entity has a particular attribute at a certain time or during a particular time interval (other than the attribute location, which is covered under relational events), for example: "Terrex 534 operates in the high density mode". A relational event involves entities which stand in an n-ary relation with each other at a certain time or during a fixed time period. The importance of the subclasses of world point and world point qualification events is in defining the world line of an entity, say the track of a ship or submarine. Of these distinctions, the most relevant for this discussion are those involving meta-events and non-meta events, and of the latter, primitive versus observational events.

In terms of the scenario described above, a primitive event may occur, say, a satellite launch by the Delta Confederation, as illustrated in Figure 4. This event, like any other event, involves the emission of energy. Such an emission is perceived by a sensing device of the Epsilon Republic. The device generates (down arrow) a report of the given event, in terms of the particular attributes of the event it is designed to measure. This sensor report is an observational event, entailing an observation of a primitive event. An S&M analyst for the Epsilon Republic accesses (up arrow) this report, which contains digitized information generated by the sensor, interprets this information as a launch event, and issues his own report about that event.

His report, which is an interpretation of the primitive event based on the observational event, is a zeroth order meta-event: the common denominator of the message traffic. At the same time, the Deltas may release an internal report about the launch, which would also constitute a zeroth order meta-event. Based on that report, the Delta press agency, NYT, may issue an announcement of the primitive event, the announcement thus constitutes a first-order meta-event. An Epsilon Republic reporter may then make an interpretation of that announcement, in the form of a report, which -- being a report of a first-order meta-event -- is therefore a second order meta-event. Corrections or other changes made by Epsilon reporters to these messages constitute a third meta level of reporting event, since they may reference reports of reports of events.

The model thus far accounts for the event reporting structure which underlies the Delta/Epsilon scenario, but we must also account for the reporter's comments about the event — i.e., his interpretation or evaluation of the event -- which can occur at any of these levels.

The reporter's goal is to identify and describe all the relevant parameters of an event (exemplified by the slots in the template for a launch event, shown in the center of Figure 5) based on the observational report produced by the sensor and any other information he may have (e.g., knowledge that a replacement of a non-functioning communications satellite is likely within a given time frame). However, if the reporter's information is incomplete or imprecise, he cannot exactly describe the parameters of an event, but will give his best interpretation of the event based on what he knows. Thus he may report a launch of "an unidentified satellite", "a probable television support satellite", "a possible CE satellite". In some cases, he may have enough information to make a comparative evaluation with launch events which have occurred in the past: "a new ESV", "the second CE satellite to be successfully orbited by the Deltas this year". Still another type of meta infor-

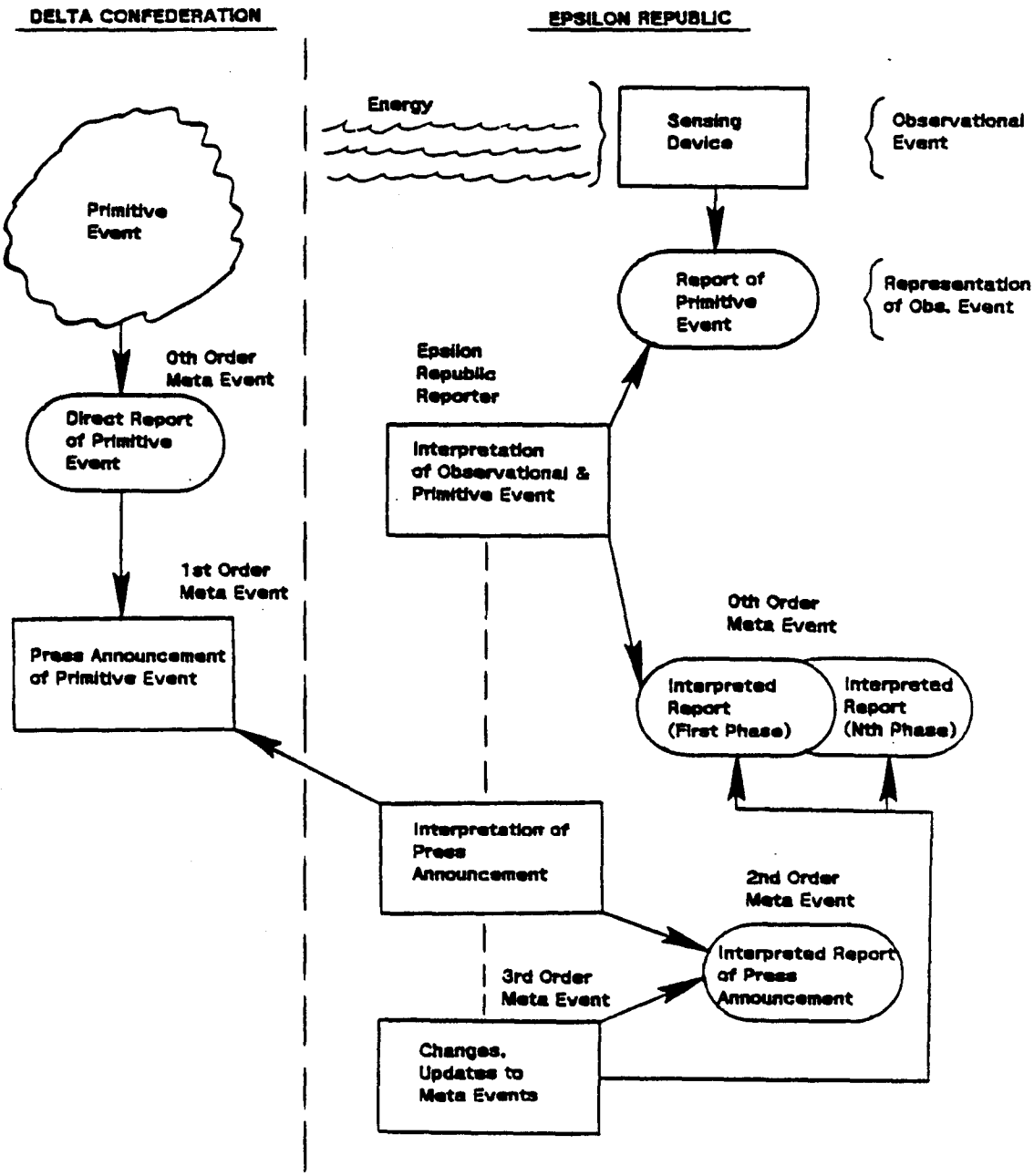

Figure 4. Model for Orders of Meta Events

mation is exemplified by the last sentence in the sample message presented above: the assignment of an identification number to a space object.

If a reporter's information is good, i.e., complete and precise, the following type of launch report is produced:

> **MSG. 04 -096**
>
> **VARIABLE DENSITY CE AGSAT LAUNCH, 20 SEPTEMBER 1983**
> **A VARIABLE DENSITY CROP ENHANCEMENT AGRICULTURAL SATELLITE WAS LAUNCHED FROM THE HARRISBURG MISSILE AND SPACE COMPLEX (HMSC) AT 1605Z ON 20 SEPTEMBER 1983. THE BR02 LAUNCH SYSTEM WAS USED TO PLACE THE SATELLITE INTO A 3 DEGREE ORBIT. WSJ OBJECT NUMBER 6018 HAS BEEN ASSIGNED TO THE PAYLOAD.**

However, when his information is imprecise and his knowledge can add little to it, he must resort to the qualified or meta-commented types of messages described above.

In order to accommodate such qualified and meta-commentary types of information, each event template may have associated with it one or more meta templates containing interpretive or evaluative information. Thus, as represented in Figure 5, an instantiated launch template produced from an observational event and a primitive event (a zeroth order reporting event, as illustrated in Figure 4) may have several additional qualifications (exemplified by, but not limited to, the meta templates illustrated in the figure). So, for example, a meta evaluative template[2] associated with a launch template expresses the Epsilon reporter's degree of belief or confidence in the launch parameters he reports: the object in the event template is believed by the Epsilon reporter to be a CE (or Crop Enhancement) satellite from the information presented in the observational report by the sensor, and from his own knowledge of past occurrences of CE satellite launches, as well as expectations of possible replacement launches, etc., during particular time intervals. All or none of the listed parameters for a launch event may be qualified in this way. Thus, in Figure 5, the Epsilon reporter believes that, to the best of his knowledge, the space object involved in the major launch event is a "probable" CE or crop enhancement satellite, and that the time of launch is "approximately" 1130Z.

Each meta template has fields which identify the source, as well as the time and date of the interpretive information. As opposed to the "infosource" parameter of these meta templates -- which shows the ultimate source of the information contained in the instantiated template -- the source is the actual originator of the message. Thus, in the case of the "designate" meta template, the "infosource" of the designation information (i.e., that the particular satellite launched from that site at that date and time has been designated a space object called "Terrex 534") is NYT, the news agency of the Delta Confederation, indicating that this information came from an NYT press announcement quoted (and interpreted) by an Epsilon reporter. This distinguishes such information from that represented by the "assign" meta template, where the Epsilon Republic reporting staff assign an identification number of their own to the satellite payload for future reference.

Another significant analytical tool of the Epsilon reporter in this scenario is the comparative evaluation, illustrated by the "compare" meta template. These comparisons involve events which have taken place before, in this case, launch events, and/or objects involved in such launches. As in the example shown in Figure 5, the comparison may specify an event involving the continuation of a satellite in an active status, where other such satellites are now inactive (implied comparison): e.g., "Terrex 534 is the only first generation crop enhancement satellite which is currently active."

An important function of meta templates is to represent predictive information: i.e., descriptions of events expected in the future, based on other events which have occurred in the past, or are currently in process. The "expect" template in Figure 5 expresses the presumable or expected parameters of mission duration, and consequently, the deorbit event which is anticipated for October 26.

To summarize, the function of the meta templates is to identify and delimit evaluative commentary, which isolates the factual information presented in most zeroth order meta event reports, and identifies information pertaining to credibility of the event occurrence, comparability with other similar entities and events, predictions of future related events, etc.

On the other hand, in addition to distinguishing the various levels of event occurrence, observation, and reporting, the function of the meta event structure illustrated in Figures 2 and 4 is to clearly demarcate the "Delta versus Epsilon" (in terms of the scenario described above) aspects of the messages. The reporters of the Epsilon Republic "assign" "Spacoid" and "WSJ" identification numbers for space object inventory purposes; the Deltas "designate" their own space objects with particular classes of object names, e.g., "Terrex 559". They "launch", "put into orbit", "deorbit", "recover", etc., while the Epsilon reporters "assess", determine "active" vs. "inactive" status, attribute satellite

---

2. The template and meta template structures shown in this figure are intended to be illustrative only: for example, the object, date/time group, and deorbit information constitute embedded templates linked to the main "launch" event template by pointers. In addition, there are several alternatives for more economical internal representation of meta template information, which are currently under review for the actual design and implementation of this information within MATRES and within the Active/Introspective Information System (a knowledge-based intelligent assistant, as mentioned above), which MATRES feeds.

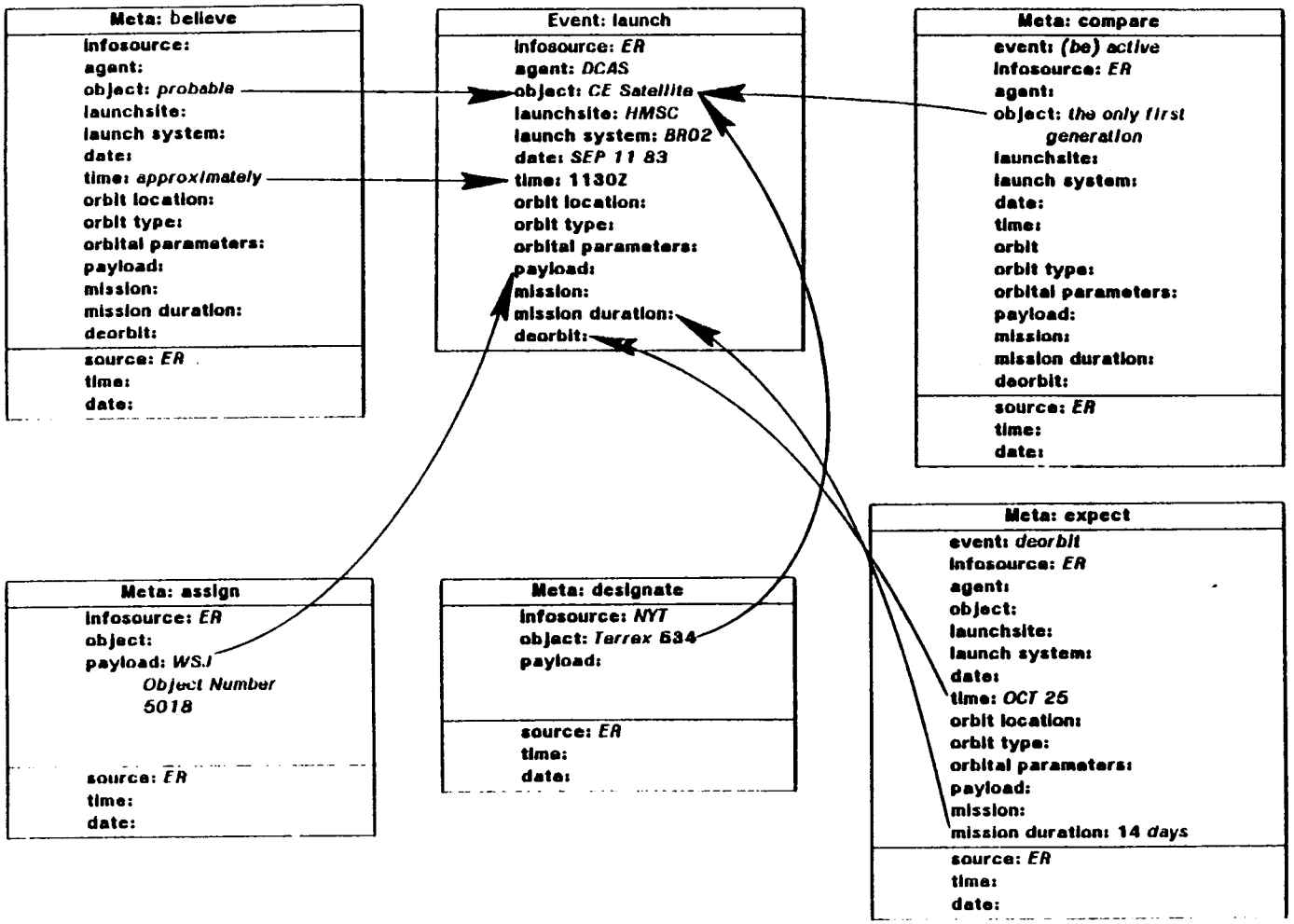

Figure 5. Representation of Meta Templates

"programs" and "maintenance" of such programs, etc., to the Deltas.

Although the detailed implementation of some of these notions remains to be worked out -- in particular, the interfaces with the Active/Introspective Information System -- we expect that the essentials of the analytical approach discussed in this paper will be demonstrable on the MATRES testbed toward the end of 1983.

REFERENCES

Dwiggins, D. L., and G. Silva [1981]
AIS Data Base Generation.
RADC-TR-81-43, Rome Air Development Center, Griffiss Air Force Base, New York, April, 1981.

Kuhns, J. L., and C. A. Montgomery [1973]
Event Record Specification System Concept: Preliminary Notions.
Operating Systems, Inc., N73-007, August, 1973.

Kuhns, J. L. [1974]
Synthesis of Inference Techniques: An Interpreted Syntax for the Logical Description of Events.
Operating Systems, Inc., N74-003, May, 1974.

Kuhns, J. L., Montgomery, C. A., and D. K. Wheichel [1975]
ERGO -- A System for Event Record Generation and Organization.
RADC-TR-75-61, Rome Air Development Center, Griffiss Air Force Base, New York, March, 1975.

Montgomery, C. A., and E. H. Ruspini [1981]
The Active Information System: A Data-Driven System for the Analysis of Imprecise Data
Proceedings of the Seventh International Conference on Very Large Data Bases, pp. 376-385, IEEE Computer Society Press, 1981.

Ruspini, Enrique H. [1982]
Possibility Theory Approaches for Advanced Information Systems.
IEEE Computer, Volume 15, Number 9: pp. 83-91, September, 1982.

Silva, G., and C. A. Montgomery [1977]
Automated I&W File Generation.
RADC-TR-77-194, Rome Air Development Center, Griffiss Air Force Base, New York, June, 1977.

Silva, G., Dwiggins, D. L., and J. L. Kuhns [1979a]
A Knowledge-Based Automated Message Understanding Methodology for an Advanced Indications System.
RADC-TR-79-133, Rome Air Development Center, Griffiss Air Force Base, New York, June, 1979.

Silva, G., Dwiggins, D. L., and C. A. Montgomery [1979b]
Satellite and Missile Data Generation for AIS.
Final Technical Report, Operating Systems, Inc., R79-037,