Some Aspects of Learning and Reorganization in an Analogical Representation

Michael Mohnhaupt, Bernd Neumann,
University of Hamburg,
Department of Computer Science,
Bodenstedtstr. 16, D-2000 Hamburg 50,
West Germany.
mohnhaupt@rz.informatik.uni-hamburg.dbp.de
neumann@rz.informatik.uni-hamburg.dbp.de

1 Introduction

This paper is concerned with learning and reorganization in an analogical representation of time-varying events. Current research in machine learning deals mainly with different aspects of acquiring knowledge for propositional representations in traditional machines (see *Michalski + Carbonell + Mitchell 83* and *Michalski + Carbonell + Mitchell 86*), or with learning problems in parallel distributed architectures (see *Rumelhart + McClelland 87*).

We discuss several learning problems with respect to analogical representations, because we believe in the usefulness of analogical representations for certain classes of problems. Theoretical considerations (see e.g. *Levesque 86*, *Sober 76*, *Palmer 78* and *Selman 87*), practical considerations (see e.g. *Larkin + Simon 87*) and experiments with the human cognitive system (see e.g. *Block 81*, *Kosslyn 80*, *Finke 85*) support this assumption.

Some aspects of categorization in our work are related to work by *Drescher 87* and *Phelps + Musgrove 85*, but our main focus is investigating learning in analogical representations.

Our representation deals with observed trajectories of moving objects. By modelling such spatio-temporal events we want to support certain tasks involving knowledge about typical object motion such as trajectory prediction. We consider this problem in the context of a natural-language guided scene analysis task where simple questions like:

Did a car turn off Schlueterstreet?

are used to initiate top-down controlled image sequence analysis. In order to control the vision processes effectively it is essential to provide knowledge about the spatio-temporal constraints implied by the verb 'turn-off' and other verbs for that matter. For a detailed discussion of the analogical representation involved in solving these problems see *Mohnhaupt + Neumann 87*, *Mohnhaupt 87*.

The main purpose of this paper is to describe processes, working on an analogical representation, to realize certain learning objectives. The problems we want to tackle are:

- learning a representation from examples,
- generalizing information from examples, abstracting relevant information from irrelevant details and adapting information to slightly different situations in the same environment,
- reorganizing the representation by computing analogies to solve similar problems in a different environment.

In Section 2 we discuss advantages and disadvantages of 'empirical' and 'analytic' representations. We believe that such a dichotomy is useful to distinguish between different knowledge representation schemes. Roughly speaking, in an 'empirical' representation, knowledge is seen as accumulated and abstracted from examples (concrete observations). On the other hand, we call a representation 'analytic', if it results from a theoretical model of the domain.

The representation, which we call 'Trajectory Accumulation Frame' (TAF), will be developed in Section 3. We show how concrete observations are recorded and how abstractions, generalizations and small adaptions are computed by local operations.

In Section 4 we describe ideas on reorganizing the representation in order to apply the representation to similar problems in a different context. If, for example the representation for a 'turning right' event was recorded in a particular geometric environment, we want it to be applicable to similar events involving streets of different shape. This reorganization can also be seen as computing analogies to make experience gained in a particular environment to be applicable to different geometric surroundings. Extracting invariant event properties is very important for this kind of reorganization. We discuss the role of perceptual primitives for the computation of invariants.

Finally, Section 5 serves to summarize the main ideas and to sketch some unsolved problems.

The representation has been implemented on a Symbolics 3600 and examples are simulated to illustrate the performance of the proposed model.

2 'Empirical' and 'Analytic' Representations

Knowledge representation schemes are often characterized along different dimensions such as 'procedural' vs. 'declarative' and 'local' vs. 'distributed'. We want to point out in this section that a distinction between 'empirical' and 'analytic' might also be useful. We call a representation 'empirical' if it is accumulated and abstracted from concrete observations, whereas we call a representation 'analytic' if it is motivated by a theoretical model of the domain. To be more concrete, we turn to a clarifying example.

Let us consider the problem of predicting the path of a thrown baseball. If one knows something about elementary physics, one should be able to calculate the path without ever having seen a throw before. By using a theoretical model which takes into account the initial forces, gravitation, the initial angle of elevation, the weight of the ball and so on, one could build a representation which calculates the path of any ball, given the necessary initial data. The relevant theoretical model on which this calculation is based is a result of a careful analysis of the problem domain. Hence we call this representation 'analytic'.

On the contrary, one could solve the prediction problem in a completely different way. By observing a certain number of instances of thrown balls, one can build a representation which relies only on concrete observations and operations operating on observed data. Generalizations and abstractions are needed in cases where the current prediction situation differs from the set of stored observations. If, for example, the initial angle of elevation differs in a prediction task from the observed and stored examples, one can use some kind of interpolation or 'blurring' (generalizing the information to slightly different situations) between the two closest examples to calculate a prediction. Knowledge is seen as accumulated and abstracted from experience in this class of representations. Hence we call the representation 'empirical'.

Comparing the two different solutions, it is important to note the following:

- A theoretical analysis and therefore a complete understanding of the domain is not needed for an empirical represention, whereas for an analytic representation a theoretical understanding is a necessary precondition. One can think of many situations where a theoretical understanding is advantageous. On the other hand, there are problems which cannot be modelled theoretically, due to an overwhelming complexity or due to the lack of a theory. In both cases an empirical solution can be the only way to tackle the problem. In cases where the time constraints imposed on the problem are very hard an empirical solution may be the only way, under the assumption that it is less complex. Some kinesthetic tasks belong to these kind of problems: We doubt that a robot can walk through an unknown environment with appropriate speed, or play tennis (besides other unsolved problems), by solving the standard kinematic and dynamic equations of robot motion. solution is more adequate.

- An empirical represention may lead to an explanation how problem solving strategies have evolved through evolution. An empirical representation has a natural transition from concrete observations to accumulated experience. There is striking evidence that empirical representations are favored by biological systems, because their representations have evolved through evolution. But empirical representations are also important for technical systems, by two reasons. First, as biological systems are often very powerful and sophisticated, typical biological solutions may inspire the developement of technical systems. And second we believe that technical systems cannot solve a certain class of problems without being able to learn and act based on experience with their environment.

- A representation should reflect the perceptual capabilities, if it models parts of the visual world. An empirical representation may implicitly lead to a natural connection between visual processes and learned behavior. In an analytic represention this connection has to be formalized explicitly. A similar argument concerns the representation of prototypes. Prototypes may easily be explained as accumulated and processed observations in an empirical representation. Again, this requires a careful modelling in an analytic representation.

We argued that it is useful to distinguish between analytic and empirical knowledge representation schemes. Criteria to evaluate how adequate a representation is for a certain task are still far from being clear in the knowledge representation literature. The term 'adequate' is used to say that the knowledge base should be tractable (*Levesque 86* use the word 'vivid' to characterize a certain form of adequate knowledge bases). Dimensions like procedural, declarative, analogical, propositional, local, distributed etc. help to distinguish between different representations, but do not imply adequacy criteria. Empirical is another useful characterization. It might turn out that a representation has to be empirical for a certain class of problems to be adequate.

3 Trajectory Accumulation Frames

In this section we describe a computational model for learning from observations in an analogical representation. We describe local operations to compute generalizations and abstractions of the accumulated experience. Finally, we show how to adapt the representation to constrained situations (e.g. obstacles in the street traffic domain). We introduce a TAF as the basic data structure for recording trajectories (i.e. object motions) in a fixed environment. Our domain of interest is street traffic, hence we will restrict ourselves to planar motion.

A TAF is a four-dimensional accumulator array C(x, y, d, v) covering a certain subfield of the xy-plane. For each xy-pair there are counter cells for all possible velocity vectors, each represented by direction d and speed v. The vector $\mathbf{S} = (x\ y\ d\ v)$ describes the motion state of an object at a given time. Note that it is composed of quantities which may be perceived by the observer of a visual scene. For each object trajectory, a trace of state vectors is registered in the TAF by incrementing the associated counters. As more objects are entered, more cells (possibly the same) are incremented without discriminating between different objects.

Let us consider now the recall of individual trajectories from a TAF. Given a starting cell the obvious operation to perform is to look for a nonzero counter in the four-dimensional vicinity. This reflects the assumption that the trajectory has been continuous and sufficiently densely sampled. Hence successive state vectors must be similar. Note that not all xy-neighbors may be reached if the velocity direction is restricted to vary smoothly. Details of the prediction algorithm can be found in *Mohnhaupt 87* and *Mohnhaupt + Neumann 87*.

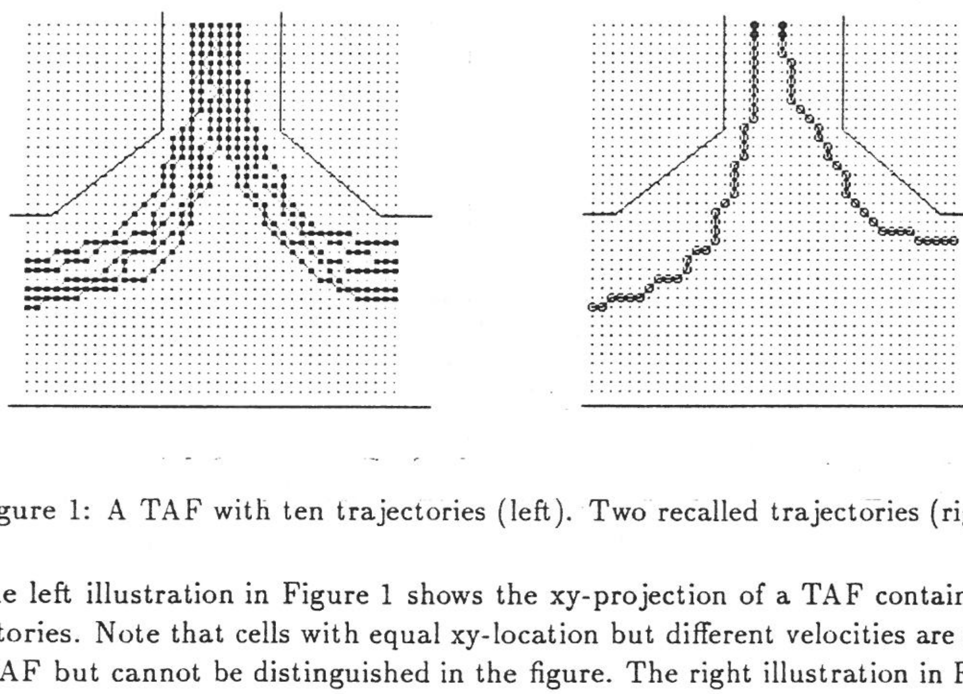

Figure 1: A TAF with ten trajectories (left). Two recalled trajectories (right)

The left illustration in Figure 1 shows the xy-projection of a TAF containing ten trajectories. Note that cells with equal xy-location but different velocities are distinct in a TAF but cannot be distinguished in the figure. The right illustration in Figure 1 illustrates the recall of two trajectories given two starting points (solid).

The number of individual trajectories that may be recalled from a single TAF depends on the similarity of the trajectories as compared with the coarseness of the quantization. Trajectories are inseparable if they meet in state space, i.e. have similar velocities at close locations. In this case the recall algorithm may continue with the 'wrong' trajectory. As we are interested in predicting typical behavior rather than individual trajectories, we do not worry about this and turn to the multiple trajectory situation.

3.1 Typical Trajectories and Prototypes

Now we consider the case where trajectory predictions are no longer determined by individual examples but depend on the experience gained from observing a possibly very large number of trajectories. We shall also introduce a blurring operation which spreads out the counter function over a local neighborhood. The combined effect of having many trajectories and of blurring will produce a counter function which is nonzero throughout the approximate area covered by the observed trajectories. Hence the TAF may be considered a four-dimensional density field with high values indicating experience supported by many observations.

In addition we shall discuss a convergence operation. This operation computes an abstraction of the current TAF by emphasizing trajectories with high probability and by suppressing trajectories with lower probability.

Traces along density maxima define a pattern of typical behavior, called the skeleton of the TAF. Predictions will essentially follow this pattern. In the remainder of this section we outline the general ideas.

Blurring:

Blurring is a generalization operation to the effect that experience represented by a counter cell is propagated to its neighbors. This is accomplished by replacing the value of each cell by the weighted average of all neighbors orthogonal to the direction of motion. Cells along the direction of motion contribute according to their positive difference.

Figure 2 shows predictions after different degrees of blurring. Note that predictions are possible at coordinates where no experience was accumulated and that predictions do not necessary follow any of the recorded individual trajectories (compare to Figure 1). The amount of blurring depends on the resolution of the representation in its different dimensions and the validity of a homogeneity criterion (blurring across the street is useful, but blurring from the street across the sidewalk should be avoided).

Convergence:

While blurring is an operation which generalizes information into the neighborhood, convergence is an operation which abstracts the most relevant information from TAFs by suppressing details and making the most important information explicit. Roughly speaking, each cell **S** adds weight to all neighboring cells from where the cell **S** can be reached, if the convergence operation is applied. The weight is proportional

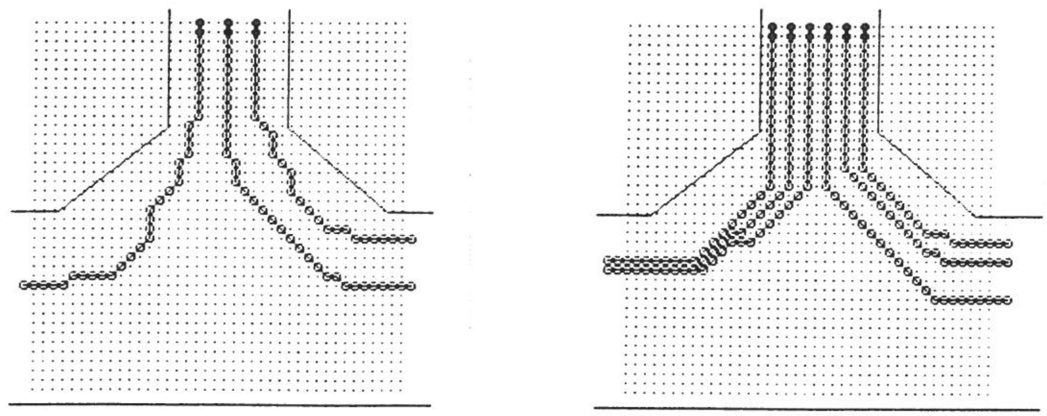

Figure 2: Predictions after single blurring (left). Predictions after triple blurring (right)

to neighboring counter values. Therefore cells with higher values get more additional support than cells with lower values.

We show the effects of the convergence operation after discussing why skeletons play an important role for our representation.

Skeletons:

Predictions are computed by picking maximum counter value cells for successors, hence cells which are local maxima play a special part. They form a pattern of typical trajectories in the sense that they outline distinct paths which are maximally supported by experience. We call this pattern a skeleton.

Figure 3 demonstrates the different effects of the blurring and the convergence operations. The skeletons show typical behaviour regardless of starting points. The left illustration in Figure 3 shows that the blurring operation has generalized the information by propagating it into the neighborhood. After applying the convergence operation in the right illustration one obtains the skeleton that makes the most important paths explicit. It is interesting to note that the skeleton can be segmented into two conceptual clusters ('turning right' and 'turning left').

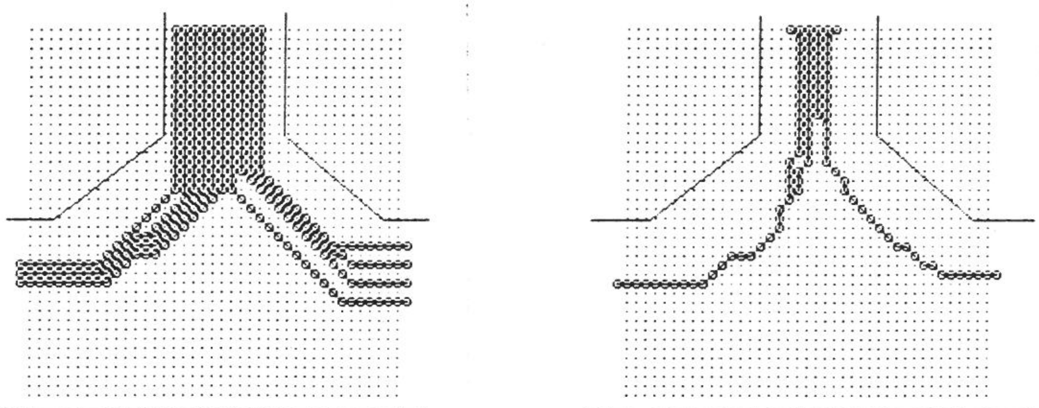

Figure 3: Skeleton of the TAF in Figure 1 (triple 'blurring'). Skeleton of the same TAF after single 'convergence' (right)

3.2 TAFs in constrained situations

Now we want to show how TAFs can be adjusted to cope with slightly different situations, e.g. situations with obstacles. We assume that obstacles were absent when the experience was accumulated. We show that a TAF is flexible enough to allow meaningful predictions in situations which are slightly different from the underlying experience.

An obstacle is a subspace of the 4-dimensional TAF where activities are not allowed, for example, due to a parking car or due to a forbidden range of velocities (in case of snow on the street). We introduce an obstacle into the TAF by setting the counter values of the appropriate subspace to zero. Then we propagate this information through the TAF by using a local inhibition operation.

Inhibition:

We inhibit a cell **S** by setting its counter value to zero under the following conditions:

- All the counter cells which can be reached from **S** are equal to zero,
- or all the counter cells from which **S** can be reached are equal to zero.

This operation is performed repeatedly for all cells until no more changes occur.

Thus one is sure that all cells which have no active predecessor or no active successor are set to zero. We demonstrate the inhibition operation with an example.

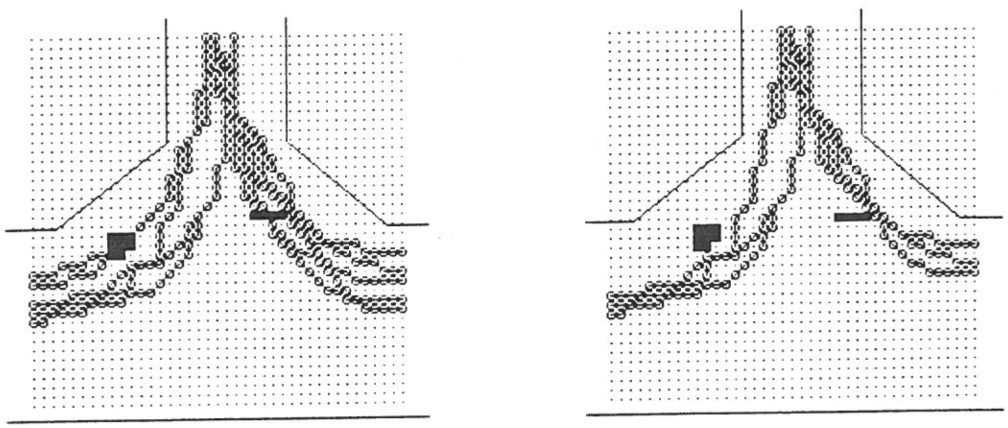

Figure 4: Skeleton for the TAF in Figure 1 with obstacles (left). The same Skeleton after inhibition (right)

The left illustration of Figure 4 shows the TAF of Figure 1 after introducing two obstacles. The effects of the inhibition operation are visible in the right illustration. All trajectories which would pass through the obstacle are inhibited. Hence the representation is adapted to the constraint situation.

4 Reorganizing the Representation

We now turn to consider how information derived from the experience aqcuired in a particular environment could be reorganized to make relevant information available when confronted with similar, albeit novel, situations. In particular, we want to make the experience accumulated at a particular street intersection applicable to another intersection with a different shape, as we cannot have direct perceptual experience for every possible geometric environment. This reorganization can also be seen as computing analogies.

4.1 Invariant event properties

The key idea is to make certain characteristic properties of an event-type explicit, where they were previously only implicit in the TAF. The descriptive properties common to all events of a given event-type are called invariant event properties. Such invariants form the necessary and sufficient conditions for the classification (categorization) of an event. The aim is to derive a particularly convenient set of descriptive primitives given the tasks of interest.

Invariant event properties are based on a subset of the rich repertoire of perceptual primitives exploited by perceptual systems in general. As such they help to identify events, to make predictions about, and to reason over, the visible world. Selecting invariant event properties can also be viewed as a first step from concrete image-like perceptions to more abstract (possibly propositional) descriptions. For example, they may contain predications about quantitative attributes of an event-type.

The set of invariant event properties should:

- facilitate tasks (be useful for the later stages of processing),
- be robustly and efficiently accessible from the data,
- provide a complete representation with respect to the tasks for which they are needed,
- be relative independent.

Some of the primitives might not be independent in a strict sense. Rather independence is defined only in terms of the information explicit (see *Levesque 86*), i.e. the information accessible with little or no computation according to the primitive operations available in the system. Therefore independence means not easily derivable within the time available. For example, we would call velocity and acceleration of an object to be independent although information about the speed of an object over time contains information about its acceleration, but only implicit.

Hence, we believe that the following features constitute a useful subset of perceptual primitives for characterizing time-varying events in terms of invariants, independent of a particular geometry:

- relative orientation between different objects,
- distance between important objects,
- relative speed between different objects,
- absolute speed,
- direction of velocity relative to a reference object,
- orientation change,
- orientation change relative to a reference object.

The repertoire of perceptual primitives also contains absolute modalities like orientation and position, which are obviously not invariant over different geometries for the time-varying events we are discussing. We choose absolute speed to be in the set of possible invariant properties, however, as we assume a stationary observer.

It is interesting to note the similarity between these primitives and the primitive events proposed for event recognition from a geometric scene description by *Neumann + Novak 86*, *Neumann 87*. They derive primitive events by analyzing natural language motion verbs, which are taken to define complex events. These primitives can be computed from a quantitative scene description. Constant values (like constant velocity or constant motions), restricted values (like 'parallel', 'close to' or 'beside'), comparative values and constant derivatives (like constant acceleration) form a basic set of primitives in their framework. Their main concern is to generate a natural language description of time-varying scenes. Therefore a propositional event description based on such primitives is a useful intermediate representation. Because our goal is to use event descriptions in an image-like context, our representations looks different. Nevertheless, both cases have strong similarities with respect to the set of perceptual primitives and the process of abstraction.

After discussing the need for useful perceptual primitives, we turn now to the different computational steps necessary to reorganize the representation.

4.2 Computing the spatial analogues

The main steps necessary to reorganize the representation to compute the spatial analogues are summarized in Figure 5:

Below, according to the four boxes in Figure 5, we discuss how the original TAF is extended, how a generic event description is computed, and how this generic description is adapted to the new environment.

- **Event model for observed environment**

  A detailed description of the original TAF was already given in Section 3. In the case of 'turning-off', the TAF contains the dimensions (x, y, d, v). The dimension were chosen because of their general usefulness for the description of time-varying scenes. Without knowing much about the events to be detected, the location and the velocity of the objects form a very general first description.

- **Extended event model**

  The extended event model contains the original dimensions (x, y, d, v), as well as all dimensions which could in principle be invariant for time-varying events. They are listed in Section 4.1.

  In our example we use only a reduced repertoire of perceptual primitives from which some are found to be invariant. For example, we do not consider the

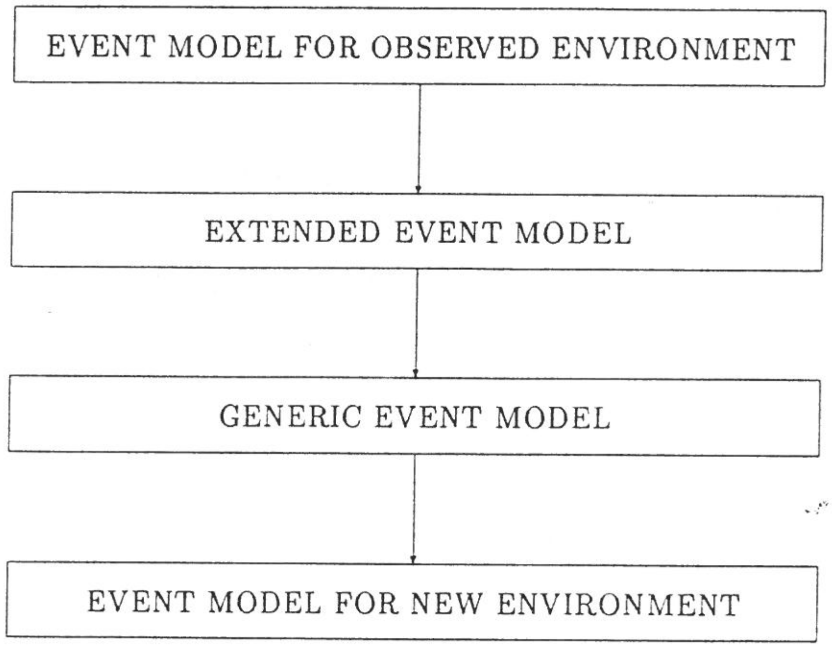

Figure 5: Transformationdiagram

orientation of an object relative to a reference object as possible invariant feature because the objects are currently represented by their respective centers of mass.

The information in the additional dimensions was already implicit in the original TAF (under the assumption that the stationary backround is known). They are now made explicit under the needs of certain tasks, involving computing generic event models.

- **Generic event model**

  The next step is to compute a generic event description by abstracting from variant information. In the case of 'turn-off' we choose as a first approximation the following dimensions to describe the invariant event properties:

  - speed (v),
  - relative orientation between car and sidewalk (ro),
  - distance between car and sidewalk (di).

Hence, the abstracted TAF for 'turn-off' contains only the dimensions (v, ro, di). The other dimensions depend on the particular geometry of the intersection.

The abstracted information can be collected from a single instance (e.g., the closest example according to some criterion) or possibly a large number of instances.

We are currently dividing invariant from non-invariant dimensions interactively. In principle this could be done automatically by observing the statistical behavior of the different dimensions over a number of observed examples, but this is left for future work.

The generic description contains data about the relative orientation between car and sidewalk and about the distance between car and sidewalk irrespective of the xy-location. This reflects the assumption that these primitives are approximately constant over the intersection for a 'turning-off' event. Therefore the generic event model stores less information than relative orientation of velocity and distance to the sidewalk for each cell (roughly speaking the values of its dimensions are 'averaged' over xy-locations). It might turn out that this assumption is too strong for a more detailed model, but it is a good first approximation.

- **Event model for new environment**

  Finally one has to adapt the generic information to a new instance (the new geometric environment) by reintroducing the dimensions (x, y, d), which depend now on the new intersection. The final TAF has the same dimensions as the original TAF. Each point in the TAF for the new intersection has a certain distance to the sidewalk and a certain orientation relative to the sidewalk. Its value is set according to the activity in the generic description for this combination of coordinates. After filling the new TAF, the local operations 'blurring', 'convergence' and 'inhibition' are applicable.

The following examples demonstrate the effects of the transformation:

The left illustration in Figure 6 shows the xy-projection of a TAF containing five trajectories. The right illustration shows the skeleton of the same TAF after single 'blurring'. The skeleton is not very narrow because no 'convergence' operation was applied. The 'blurring' operation spreads the information in areas where no experience was accumulated (see Section 3).

The left illustration of Figure 7 shows the skeleton of the TAF obtained from the street shape in Figure 6. This is the main result in this section. Information collected in a particular environment in space-time has been transformed to be applicable in a different environment by conserving event dependent information and abstracting away irrelevant information. This figure can also be interpreted as showing the computed visualization of 'turning-right' events on intersection 2 in analogy to observed 'turning-right' events on intersection 1.

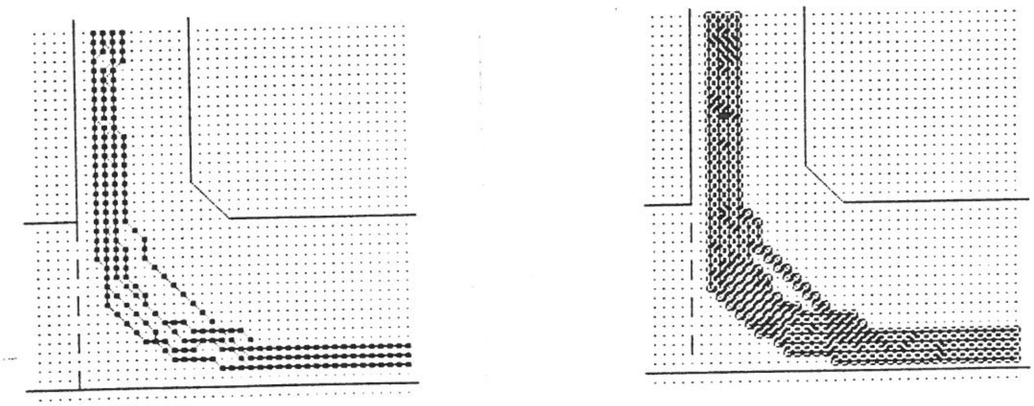

Figure 6: A TAF with five trajectories (left). Skeleton of this TAF after single 'blurring' and no 'convergence' (right)

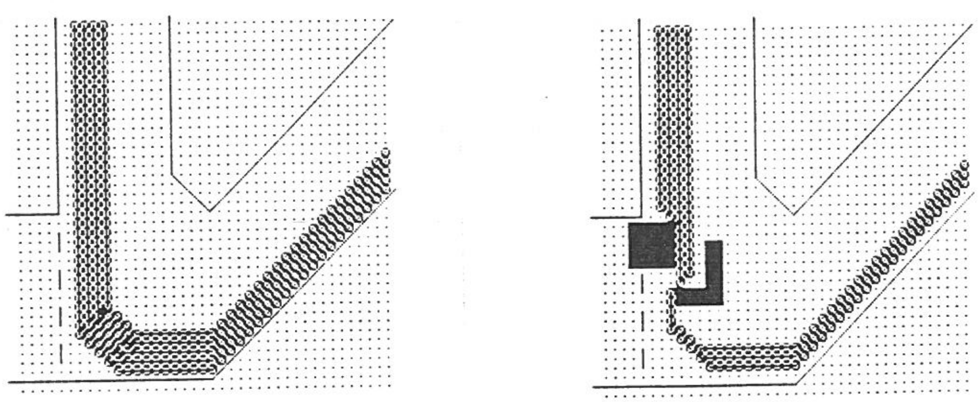

Figure 7: Skeleton of the transformed TAF (left). Skeleton after introducing obstacles and 'inhibition' (right)

The analogy is computed by using invariant event properties which were choosen from a rich set of perceptual primitives.

The right illustration in Figure 7 demonstrates the 'inhibition' operation after introducing obstacles in the transformed TAF. The local operations are applicable to image-like representations like TAFs, regardless whether they are a result of direct observation or a reorganisation for a particular purpose.

5 Summary and Discussion

We have considered several aspects of learning and reorganisation in analogical representations. We believe that learning, reorganizing knowledge, and having an analogical representation are essential to build an adequate representation for a certain class of perceptual and cognitive tasks.

We have introduced the dichotomy of 'empirical' vs. 'analytic' as a useful charaterization of a knowledge representation scheme. A representation is empirical if the knowledge is seen as abstracted and accumulated from concrete observations, whereas a representation is analytic if it is the result of a theoretical model of the domain. An empirical representation is advantageous in different situations. It provides for a natural transition from individual examples to accumulated experience. There is no epistemological discontinuity between concrete examples and prototypical knowledge. In addition, it leads to a possible explanation how a problem solving strategy has evolved through evolution. Furthermore, there are problems which cannot be modelled in other than an empirical way, due to a lack of understanding, due to an inherent complexity or due to constraining time conditions.

A TAF and its associated local operations have been proposed as an empirical computational model for:

- learning an analogical representation from concrete observations,
- generalizing information from examples, abstracting relevant information from irrelevant details and adapting information to slightly different situations
- computing analogies by reorganizing the representation to solve similar problems in a different context.

The model has the following interesting properties:

First, in a TAF one can distinguish between typical and atypical behavior, thus it captures an important dimension of experience. In the context of top-down controlled image sequence analysis, for example, high-value areas of a TAF can be used to define primary search areas for - say - discovering a turn-off event.

Second, a TAF can be used for generalizing observations w.r.t. location and velocity. These are essential means to help apply experience to novel but similar situations.

TAFs are also able to deal with obstacles which were not present during knowledge acquisition.

Third, a TAF may be characterized by abstracted information (its skeleton) and by the trajectory prototypes as defined by the skeleton. This provides the possibility to refer to and reason about typical motion behavior in terms of a finite set of alternatives.

Fourth, TAFs can be used to compute analogies by reorganizing the representation appropriately. 'Turn-off' events have been predicted for an intersection where no examples have been observed. The prediction has been computed in analogy to 'turn-off' events observed at an intersection with different geometry. Hence an important aspect of reasoning about experience has been modelled.

Fifth, the transition from event descriptions gained in particular environments to generic event models was achieved by using a subset of perceptual primitives to characterize invariant information. We feel that the set of perceptual primitives we have discussed plays an important role for several perceptual and cognitive tasks.

The flexibility of the representation was achieved by using an analogical, 'scenic' mapping from the world into the model and by applying local operations to make certain information explicit. We believe that an analogical representation is needed for efficient solutions in the domain discussed in this work.

Several aspects of the problem area will be investigated in the near future. One is the relation between analogical image-like representations and more abstract propositional representations. How can we learn propositional descriptions from observations? Are image-like representations a useful intermediate representation for the computation of propositions?

Another problem concerns the limits of separability. How can one attain the separation of trajectory bundles which share common parts? One way to deal with this problem is to extend the state vector **S** to contain other useful features such as object characteristics or - to improve separation - global trajectory characteristics. As the usefulness of such features would depend on the situation, a theory of how to select useful features must be developed. This, again, is left for future work.

A third problem we want to tackle was already mentioned in Section 4. We want to distinguish variant from invariant information automatically by looking at the distribution of information in different dimensions over a variety of observed examples.

Finally, the role of different characteristics of representations (like empirical and analytic) is not completely understood. Can we develop systematic criteria to evaluate the adequacy of a certain representation?

Acknowledgements

This work has been financially supported by DFG (German Science Foundation). We thank David Fleet for valuable comments on Section 4.

References

[Block 81]
Imagery.
Ned Block(Ed.).
MIT Press, Cambridge, Mass. 1981.

[Drescher 87]
A Mechanism for Early Piagetian Learning.
Gary L. Drescher.
Proc. Nat. Conf. on Art. Intell. AAAI-6, 1987, 290-294.

[Finke 85]
Theories Relating Mental Imagery to Perception.
Ronald A. Finke.
Psychological Bulletin **98** (1985) 236-259.

[Kosslyn 80]
Image and Mind.
Stephen M. Kosslyn.
Harvard University Press, 1980.

[Larkin + Simon 87]
Why a Diagramm is (Sometimes) Worth Ten Thousand Words.
Jill H. Larkin, Herbert A. Simon.
Cognitive Science **11** (1987) 65-99.

[Levesque 86]
Making Believers out of Computers.
Hector J. Levesque.
Artificial Intelligence **30** (1986) 81-108.

[Michalski + Carbonell + Mitchell 83]
Machine Learning I.
Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell.
Tioga Publishing Company 1983.

[Michalski + Carbonell + Mitchell 86]
Machine Learning II.
Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell.
Morgan Kaufman Publisher 1986.

[Mohnhaupt 87]
On Modelling Events with an Analogical Representation.

Michael Mohnhaupt.
Proc. German Workshop on Artificial Intelligence GWAI–11, 1987, 31-40.

[Mohnhaupt + Neumann 87]
Szenenhafte Modelle fuer zeitabhaengige Ereignisse.
Michael Mohnhaupt, Bernd Neumann.
Technical Report, FBI-B-127, Fachbereich Informatik, University of Hamburg, February 1987.

[Neumann 87]
Propositionale und analoge Repraesentation von Ereignissen.
Bernd Neumann.
Mitteilung Nr. 149, Fachbereich Informatik, Universitaet Hamburg 1987.

[Neumann + Novak 86]
NAOS: Ein System zur natuerlichsprachlichen Beschreibung zeitveraenderlicher Szenen.
Bernd Neumann, Hans-J. Novak.
Informatik Forsch. Entw. 1 (1986) 83-92.

[Palmer 78]
Fundamental Aspects of Cognitive Representation.
S. P. Palmer.
in: E. Rosch, B.B. Lloyd (Ed.): Cognition and Categorisation, Hillsdale, N.Y.: Erlbaum Press 1987.

[Phelps + Musgrove 85]
A prototypical approach to machine learning.
R. I. Phelps, P. B. Musgrove.
Proc. Int. Joint Conf. on Art. Intell. IJCAI–9, 1985, 698-700.

[Rumelhart + McClelland 87]
Parallel Distributed Processing I + II.
David E. Rumelhart, James L. McClelland and the PDP research group.
MIT Press, Cambridge, Mass. 1987.

[Selman 87]
Analogues.
Bart Selman.
Technical Note CSRI-47, University of Toronto, Department of Computer Science, March 1987.

[Sober 76]
Mental Representations.

Elliot Sober.
Synthese **33** (1976) 101-148.