Artificial Intelligence Review **10**: 235–252, 1996.
© 1996 *Kluwer Academic Publishers. Printed in the Netherlands.*

Automatic Reconstruction of Vasculature

PAUL MC KEVITT

Department of Computer Science, Regent Court, University of Sheffield, 211 Portobello Street, GB- S1 4DP, Sheffield, England, U.K., E.U.; E-mail: p.mckevitt@dcs.sheffield.ac.uk

and

PETER HALL

Department of Computer Science, P.O. Box 600, Victoria University Wellington, Wellington, New Zealand; E-mail: peter@comp.vuw.ac.nz

Abstract. Two of the most difficult problems in Artificial Intelligence are processing visual scenes and processing natural languages. There has been a large amount of research in each of these fields but little on their integration. This is surprising given the potential importance of integrated systems, not only for understanding human cognition but also for the range of practical applications that will be enabled. We review previous work and provide an overview of our own work. We focus upon the medical application of reconstructing complicated cerebral blood vessel structures and associated pathologies from images and medical reports. This gives our work a clear and significant practical aim. We show how the ostensibly disparate technologies can be married using a single knowledge representation. Previous attempts at reconstruction have used images alone and no satisfactory solution exists. We believe that the synergy provided by integrating vision and natural language processing provides an information-rich environment that will enable progress toward an efficient and robust solution. Such an integration will have not only have important practical uses but also implications for Artificial Intelligence, Cognitive science, Philosophy, and Psychology.

Key words: angiogram, artificial intelligence, natural language processing, text processing, vasculature, vision processing.

1. INTRODUCTION

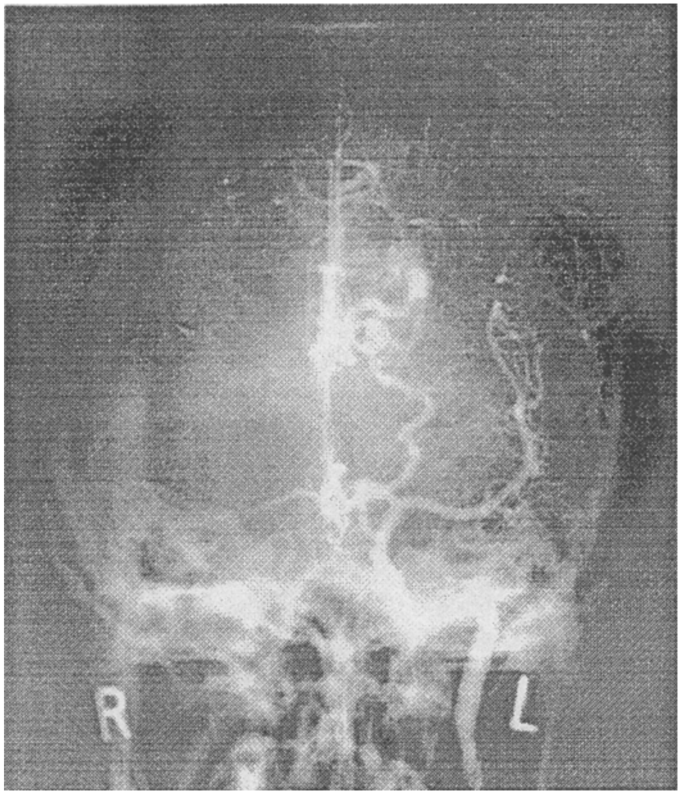
Humans are able to combine the processing of vision and language, apparently with ease. For instance, humans can use words to describe a picture and can reproduce a picture from a language description. Moreover, humans can exhibit this kind of behaviour over a very wide range of input pictures and language descriptions. Even more impressive is the fact that humans can look at images and describe not just the picture itself but a set of emotions evoked by it. Although there are theories of how we process vision and language there are few theories about how such processing is integrated. There have been large debates in Philosophy and Psychology with respect to the degree with which people store knowledge as propositions or pictures (see Pylyshyn 1973; Kosslyn and Pomerantz 1977).

There has been much research in Artificial Intelligence (AI) on the processing of natural language like English and on the processing of visual scenes (see Ballard and Brown 1982; Partridge 1991). However, there has been little work on linking natural language processing (NLP) and vision processing (VP). There are at least two advantages of linking the processing of natural languages to the processing of visual scenes. First, investigations into the nature of human cognition may benefit. Such investigations are being conducted in the fields of Psychology, Cognitive Science, and Philosophy. Computer implementations of integrated VP and NLP can she light on how people do it. Second, there are advantages for real-world applications. The combinations of two powerful technologies promises new applications that include: automatic production of text from images; automatic production of images from text; and the automatic interpretation of images with text. Research into any of these must be regarded as AI, but other disciplines may benefit. For example, the first of these application areas would benefit the discipline of *scientific visualization* and its applications, the second would benefit *computer graphics* and its applications, the third would benefit AI itself, specifically in *machine learning* and *information acquisition*. The theoretical and practical advantages of linking natural language and vision processing have also been described in Wahlster (1988).

We believe that the way forward for developing general theories of language and vision processing is to focus on specific applications such as the medical domain. We are building a system that processes x-ray projections (*angiograms*) and their associated medical reports. The angiograms are separated by an angle of about ninety degrees making them *biplane angiograms*. The medical reports are prepared by expert radiologists as they examine the angiograms and describe the appearance of the vasculature in them. A typical pair of angiograms are shown in Figure 1, and an associated medical report in Figure 2. Each of these input datum relate to blood vessel structures (*vasculature*) and arteriovenous malformations (AVMs) within the human body. AVMs are congenital abnormalities of the vasculature. These AVMs are dangerous because if they hemorrhage the results can be fatal. The clinical reasons for acquiring angiograms are fully explained in Hall *et al.* (1995b).

Our goal is to reconstruct vasculature and AVMs in three-dimensional space given biplane angiograms and associated medical reports. Previously this problem has been addressed using images alone and work has concentrated on reconstructing *coronary vasculature* around the heart. The *cerebral vasculature* comprises many vessels which are smaller than the coronary vessels; and there are wide variations in the branching structure between individuals. So reconstruction of cerebral vasculature is much harder than of coronary vasculature. Our work differs in two important ways: (1) it provides a unique representation of a *collection of vasculature* and, (2) it uses information acquired from medical texts in addition to images. We believe that our approach provides an information-rich environment that will enable reconstruction of vasculature in the brain in an efficient and robust manner.

FRONT



SIDE

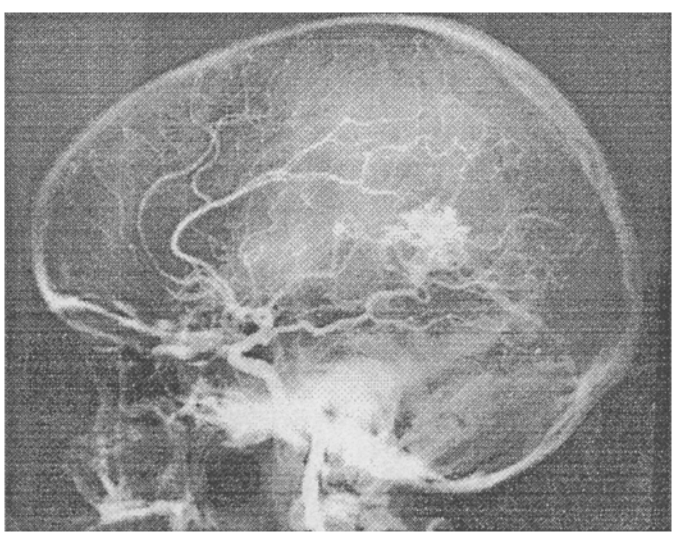


Fig. 1. An example angiogram pair.

RIGHT AND LEFT INTERNAL AND EXTERNAL CAROTID ANGIOGRAPHY, RIGHT VERTEBRAL ANGIOGRAPHY.

On the (L) side there is an arterio-venous malformation with small tortuous vessels in a nidus of a dia. of about 2cms, situated close to the midline in the forceps major, supplied mainly by two posterior branches of the occipital division of the widened (L) posterior cerebral artery, which in turn is supplied both from the carotid syphon via a wide posterior communicating artery, and also from the basilar artery. The upper part of the lesion is also to some extent supplied from two dilated small arterial branches of the posterior pericallosal artery on the (L) side. The main drainage is via the medial atrial vein to the vein of Galen and the straight sinus. Otherwise normal, free communication across all parts of the circle of Willis with the possible exception of the (R) posterior medial portion, the (R) posterior cerebral artery filling directly from the (R) carotid syphon and not at all from the basilar artery.

Fig. 2. An example medical report.

In this paper we first give the background necessary to understand our problem and discuss previous work on medical image and report analysis, and the integration of NLP and VP (see section 2). Next, we present our contribution. The knowledge representation which is a representation of a collection of vasculature is described in section 3.1. We then show how this knowledge is used by the VP module in section 3.2 and the NLP module in section 3.3. We outline an architecture that integrates the knowledge representation with both the VP and NLP modules in section 3.4. In that section we also discuss the implications and benefits of integration. We conclude in section 4.

2. BACKGROUND

Reconstruction of vasculature and AVMs given x-ray angiograms is impossible unless some information prior to the angiograms is used. This can be seen when it is realised that reconstruction from angiograms may be likened to recovering a matrix from its row sums and column sums; the problem is *under-determined.* Traditional techniques that were developed in the computer vision literature deal with images that are separated by an angle of just a few degrees (for example, see Marr 1982; Mayhew and Frisby 1981) and cannot be used here. Prior information is required so that the ostensibly disparate images can be matched. We begin by describing existing systems for reconstruction of vasculature from angiograms, systems that process medical reports, and systems that integrate NLP and VP.

2.1. *Reconstruction of vasculature from angiograms*

There are many problems associated with reconstructing vasculature and AVMs from x-ray angiograms (see Hall 1993a). Previously we have considered the problem of reconstructing AVMs from angiograms (see Hall *et al.* 1993). Here, we confine our discussion to reconstructing vasculature. A full review of systems is available in Hall (1993b) and now we describe a representative subset of systems.

Stansfield (1986) describes an early system that recognises individual vessels within angiograms. Her system, called ANGY, is composed of three major parts: a low-level image processing model for segmenting images; a mid-level module that combines segments into large-scale cohesive units; and a high-level module that recognises these large-scale units and labels them according to standard clinical terminology. Suetens *et al.* (1987) follow an almost identical line as do Rake and Smith (1987) though the latter authors use a blackboard architecture. Delaere *et al.* (1990) not only label features in angiograms with clinical names but use a Boolean expression to specify those sets of labels that are self consistent. Some control over variance in vascular topology is gained this way. Garreau *et al.* (1991) use a three-dimensional model that contains qualitative rather than quantitative descriptions. A Prolog program provides rules that drive reconstruction.

The common characteristic of most of these attempts is that they tend to be very domain specific. They describe angiograms of the coronary vasculature from specific views. In short, they use rules to describe angiograms. Rules seem ill-suited for reconstruction of the cerebral vasculature. This is principally because the wide variation of branching patterns observed in cerebral vasculature between individuals makes it a complex structure not easily described by rules. Garreau *et al.* (1991) provide an exception in that their model is three dimensional. However, their model does not appear to make anatomy explicit.

Our representation differs because (1) it explicitly represents physical vasculature in three-dimensions and is therefore less domain specific, and (2) is an adaptive representation that can learn. Not all vascular branching patterns are documented, and it is unclear that they can be, so an adaptive model is essential. In short, our model of a collection of vasculature is an information repository that is designed to handle the level of complexity manifested by the cerebral vasculature. So far as we are aware our representation is unique.

2.2. *Medical report processing*

Schröder (1992a, 1992b, 1992c) has built a natural language system called METEXA (MEdical TEXt Analysis) that can process radiological reports that are spoken. METEXA is used for analysis of such reports. Another goal of the project is to conduct continuous speech recognition. METEXA centres on a solution to speech analysis by the use of domain-specific knowledge and contains a lexicon of about 1000 forms that specify the syntax and semantics of words. METEXA has 100 grammar rules and semantic analysis is conducted by constructing a conceptual graph (see Sowa 1984). Knowledge about the domain is encoded in canonical graphs of anatomical objects, pathological alterations, radiological terms, and other more general concepts. The knowledge-base contains about 400 concept types. A rule-base is used to answer questions about radiological reports. Plan (see Wilensky 1983) and script (see Schank and Abelson 1977) knowledge structures are used to generate expectations about the next incoming utterance. These structures contain knowledge about typical contents of radiological reports.

With respect to domain knowledge there is a lattice of concept types and relations are defined between them. Concept types important for radiological applications are ANAT (anatomical), PATHO ALT (pathological), and ATTRIBUT (anatomical, pathological and radiological attributes). The definitions of valid connections or relations between concept types and relations are provided by canonical graphs (see sowa 1984, p. 91). The definitions express restrictions and some relations used are shown in Table I below.

Table I. METEXA concepts and their relationships.

concept type	relation type	concept type
[ANAT]	PATHO	[PATHO ALT]
[ANAT]	SIDE	[SIDE]
[ANAT]	LOC	[ANATLOC]
[ANAT]	AATTR	[AATTRIBUTE]

Schröder (1992c) points out that the language style found in radiological reports shows characteristics of being a sublanguage. It is also important that such sublanguages have a restricted and simplified structure compared to full natural languages. Reports are written in *telegraphic* style where standard phrases are used and where verbs may be omitted. The semantics or meaning of utterances is more important than their structure or syntax. Utterances often consist of constituents without syntactic glue. Specialised domain knowledge is important for understanding and producing such clinical reports. Certain anatomical locations and their characteristics are expected to occur in order to remove ambiguity of meaning. Although Schröder has considered the possibility of applying NLP to medical reports there has been no proposal to link this work to the processing of visual data.

Other work on NLP of clinical reports includes the Linguistic String Project (see Sager *et al.* 1987), a system for acquiring medical facts from radiological reports (see Ranum 1988), knowledge-based processing of radiological reports (see Baud *et al.* 1991), and a dialogue system for querying and updating medical databases (see Mery *et al.* 1987). Scherrer *et al.* (1989) present work on knowledge acquisition and a compilation of other work is found in Kingsland (1989). Again, none of this work is linked to any vision processing. In contrast, we apply NLP to medical reports to assist our VP module in locating AVMs and vascular reconstruction.

2.3. *Integrated language and vision processing*

A number of natural language systems for the description of image sequences have been developed (see Herzog and Retz-Schmidt 1990; Neumann and Novak 1986). These systems verbalise the behaviour of human agents in image sequences about football and describe the spatio-temporal properties of the behaviour observed. Retz-Schmidt (1991) and Retz-Schmidt and Tetzlaff (1991) describe an approach that yields plan hypotheses about intentional entities from spatio-temporal information about agents. The results can be verbalised in natural language. The system called REPLATI-II takes observations from image sequences as input. Moving objects from two-dimensional image sequences have been extracted by a vision system (see Herzog *et al.* 1989) and spatio-temporal entities (spatial relations and events) have been recognised by an event-recognition system. A focussing process selects interesting agents to be concentrated on during a plan-recognition process. Plan recognition provides a basis for intention recognition and plan-failure analysis. Each recognised intentional entity is described in natural language. A system called SOCCER (see André *et al.* 1988; Herzog *et al.* 1989) verbalised real-world image sequences of soccer games in natural language and REPLAI-II extends the range of capabilities of SOCCER. Here, NLP is used more for annotation through text generation whereas we are interested in analysis.

Maaß *et al.* (1993) describe a system, called VITRE GUIDE, that generates mutlimodal route descriptions for computer assisted vehicle navigation. Information is presented in natural language, maps and perspective views. Three classes of spatial relations are described for natural language references: (1) topological relations (e.g. "in,", "near"), (2) directional relations (e.g. "left," "right") and (3) path relations (e.g. "along," "past"). The output for all presentation modes relies on one common three-dimensional model of the domain. Again, VITRA emphasises annotation through generation of text, rather than analysis, and the vision module considers interrogation of a database of digitized road and city maps rather than vision analysis.

3. VASCULAR RECONSTRUCTION

There have been efforts devoted to reconstructing vasculature from angiograms, analysis of medical texts, and the integration of vision and language processing. This section describes our contribution to all three areas. Because our representation of a collection of individual vasculature is central to our approach we begin with that. We then describe VP and NLP processing separately, and finally bring all of these together with the description of a single architecture.

3.1. *Knowledge representation*

Our model of a collection of vasculature is the *knowledge representation* (KR) used by all other modules in our overall system. We have diverted from all previous representations of vasculature by providing a KR that models many vasculature, that models the physical anatomy of individual vasculature, that is capable of adaptive modification, and that can be used for many purposes. Indeed, we have succeeded in using our KR to simulate the complete angiographic procedure which involves animating blood flow through the vasculature (see Hall 1994a, 1994b, 1994c). Our KR partitions the vasculature into three parts: topological (branching structure), geometric (such as shapes of vessels), statistical (such as variances in shape or the frequency at which a given vessel appears). Our KR was influenced by *ModelVisual* which is a program for visualisation developed by Kunii and Shinagawa (1991). Hall and McGregor (1993) give a full description of our KR. An overview is given here.

3.1.1. *Topology*

In our model the topology of any given individual vasculature is represented by a labelled graph:

$$G_i = (V_i, E_i),$$

where V_i is a set of vertices and E_i is a set of edges. The vertices correspond to vascular points that are branches (furcations) and vessel ends. The edges are vessel segments that connect these points. A vessel, as defined in clinical texts (see Salamon and Huang 1976) corresponds to a path through our total graph. The vertices and edges are labelled with geometric and statistical information as discussed below.

A collection, L, of such graphs, each representing a distinct individual, constitutes the supposed initial state of the knowledge base and we write:

$$L = \{G_i\}.$$

The individual graphs are 'added' together to generate a *proto-graph*:

$$G^+ = \cup\, G_i\ (\cup\, V_i,\ \cup\, E_i) = (V^+, E^+),$$

where $\cup$ runs over all G_i, and in which the individual graphs appear as subgraphs. Notice that union distributes over the individual sets that comprise the graph. This is a valid step provided that the symbols are all drawn from the same alphabet, say $\mathcal{P}$ for the vertices and $\mathcal{P}^2$ for the edges. We note two things: (1) the union can be performed incrementally which is the basis of the adaptive modelling and (2) to generate such an alphabet between any pair of graphs is equivalent to solving a maximal common sub-graph (MCS) problem where the alphabet symbols are correspondences between nodes. In the general case the problem is NP- complete which makes the MCS problem intractable and our method untenable. Read and Corneil (1977) give an excellent discussion on such problems. We note that solving the MCS problem is made feasible because the graphs are richly labelled so we are dealing with a special rather than the general case.

The proto-graph is insufficient as a KR because there is no way to distinguish between its subgraphs that are elements of L and sub-graphs that are not. We wish to make this distinction because we cannot be sure that an arbitrary sub-graph of the proto-graph represents a valid vasculature. To overcome this problem we might label the nodes and arcs of the proto-graph with references that identify which graphs they came from and in many cases there will be more than one such graph. Although these labels are a sufficient representation we choose instead to use a matrix in which columns are labels of the proto-graph and the rows are valid sub-graphs (elements of L) of the proto-graph. Let's consider an example.

Consider the vasculature depicted in Figure 3, which shows both geometric and topological aspects. The following sub-structures can identified

$$G_1 = (\{a, b, c, d, e, g\}, (ab, bc, bd, ce, dg\})$$
$$G_2 = (\{f, i, j, m, o\}, (fi, fj, fm, mo\})$$
$$G_3 = (\{h, k, l, n, p\}, \{hk, hl, hn, np\})$$

To form the proto-graph for these three graphs we construct:

$$\begin{aligned} G^+ &= G_2 \cup G_2 \cup G_3 \\ &= (\{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p\}, \\ &\quad \{ab, bc, bd, ce, dg, fi, fj, fm, hk, hl, hn, mo, np\}) \end{aligned}$$

and a matrix as shown below in Table II.

Table II. Matrix representation of the example.

G^+	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	ab	bc	bd	ce	dg	fi	fj	fm	hk	hl	hn	mo	np
G_1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
G_2	0	0	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	1	1	1	0	0	0	1	0
G_3	0	0	0	0	0	0	0	1	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1

Note that columns of the matrix are equivalent to the labels on the nodes and arcs of the protograph. We can use this matrix to generate a Boolean valued expression by reading rows:

$$\begin{aligned} &\mathrm{B}\ (G^+, G) = \\ &\quad (ab \wedge bc \wedge bd \wedge ce \wedge dg \wedge \overline{fi} \wedge \overline{fj} \wedge \overline{fm} \wedge \overline{hk} \wedge \overline{hl} \wedge \overline{hn} \wedge \overline{mo} \wedge \overline{np}) \vee \\ &\quad (\overline{ab} \wedge \overline{bc} \wedge \overline{bd} \wedge \overline{ce} \wedge \overline{dg} \wedge fi \wedge fj \wedge fm \wedge \overline{hk} \wedge \overline{hl} \wedge \overline{hn} \wedge mo \wedge \overline{np}) \vee \\ &\quad (\overline{ab} \wedge \overline{bc} \wedge \overline{bd} \wedge \overline{ce} \wedge \overline{dg} \wedge \overline{fi} \wedge \overline{fj} \wedge \overline{fm} \wedge hk \wedge hl \wedge hn \wedge \overline{mo} \wedge np) \end{aligned}$$

which can be reduced to:

$$\begin{aligned} B(G^+, G) = ((&ab \wedge bc \wedge bd \wedge ce \wedge dg) \\ &\neq (fi \wedge fj \wedge fm \wedge mo) \neq (hk \wedge hl \wedge hn \wedge np)) \end{aligned}$$

where $\neq$ is the exclusive OR (XOR) proposition. We call $B(G^+,G)$ the *discrimination proposition* because it discriminates between valid and invalid sub-graphs of the proto-graph. Notice that if we wish to add a new graph to the KR then we need to update both the proto-graph and the discrimination proposition. This addition may be regarded as form of *machine learning*.

This example serves to illustrate the essential features of our KR. We have carefully defined the meaning of the matrix and derived the discrimination

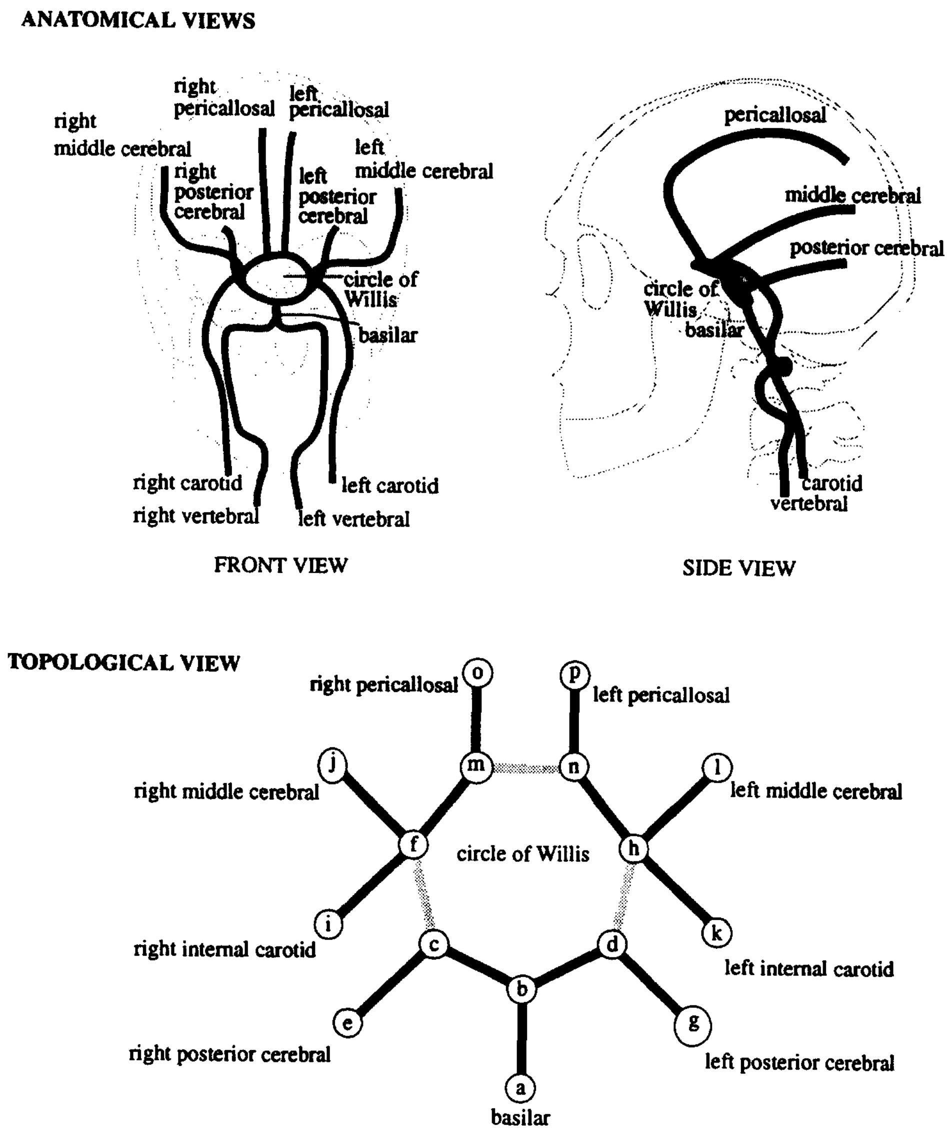


Fig. 3. A proto-typical cerebral vasculature showing the circle of Willis.

proposition from it (see Hall and McGregor 1993). We have also constructed the discrimination proposition by considering the need to recover valid sub-graphs as an injection from the proto-graph to the set all possible sub-graphs.

3.1.2. *Geometry and statistics*

Geometrical information is held in labels associated with graph features. These labels may be retrieved from the feature symbols. Labels are used to fix the three-dimensional shape of the vasculature. We define the *geometric model*:

$$F(G) = (F_V(V), F_E(E))$$

where F_V and F_E are bijective functions. The set $F_V(V)$ contains geometric descriptions of furcations and $F_E(E)$ are geometric descriptions of vessel fragments. Typically, elements of both these sets comprise generalised cylinders (see Ballard and Brown 1982). Hall (1994a, 1994b, 1994c) gives a full account of geometry.

Statistical information is also held in feature labels and we define the *statistical model:*

$$S(G) = (S_V(V), S_E(E))$$

where S_V and S_E are bijective functions. Statistical information is used to allow the shapes of vessels and furcations to vary and may also count the frequency with which a particular geometric element is observed to occur in a given population.

3.2. *The vision module*

The VP module uses the KR to enable it to predict what vessels will appear in an angiogram and to recognise them should they appear. Because the KR is three-dimensional this prediction and recognition can occur from any viewpoint. In short, reconstruction is driven on a *hypothesise and test* basis. In particular, vascular reconstruction can be considered as an attempt to prove that some arbitrary graph, H, is a valid sub-graph of the proto-graph. Hence, reconstruction can be regarded as graph matching. The description we give here is expanded upon in Hall *et al.* (1995a).

We use a search algorithm for the match. Nodes in the search space are propositions that a particular vessel fragments in the KR corresponds to particular vessel fragments that have been segmented from the angiograms from different points of view. For example,

CORR(AB, *ab*, *a'b'*)

is a proposition that vessel AB in the KR corresponds to vessel *ab* in one angiogram and *a'b'* in another. The value of this correspondence is taken to be the minimum of the individual correspondences. This is determined by geometrically matching the projected model vessel, AB, to the relevant vessel in the angiogram. We note that this match requires the vessel being matched to be oriented in space so that a plausible projection occurs. We assume that there is some heuristic available that enables a root node to be decided, otherwise we will appeal to the expert user to determine the correspondence. Given this initial correspondence a 'root' vessel can be oriented; and this gives an initial orientation to the whole KR.

The expansion of the search is controlled by the KR. The KR is consulted so that an hypothesis of which vessels should appear next in the angiograms can be generated. Vessels that are predicted to appear simultaneously are combined into a single proposition headed by AND. For example, AND(CORR(BC, *bc*, *b'c'*), CORR(BD, *bd*, *b'd'*) would indicate that vessels BC and BD in the model must simultaneously correspond to vessels bc and *b'c'*, and bd and *b'd'* respectively, in an angiogram pair. Such constructions are used where the model furcates. Where the model offers a choice between possibilities that were component to the proto-graph we use NAND rather than XOR because this admits the possibility of no component of the KR being present in the angiograms.

As the search space expands the raw correspondences are pushed to the leaves of the search space leaving propositioned operators at the branches. At each step in the expansion vessels in the KR are locally re-oriented so that the vessels that are predicted to appear have a high chance of matching a counter-part in the angiogram. The hypothesis can then be tested against the real angiograms by geometric matching, as explained above.

At any stage in the search the tree can be parsed to reveal which components of the KR currently best account for the angiograms. Those components that fail to meet a threshold can be disregarded in future expansions and this prunes the search space. The expansion stops whenever the KR or the angiograms are exhausted of data. In the ideal case this exhaustion occurs in both at the same time. The final reconstructed vasculatures are determined by a simple parse of the search tree.

3.3. *The natural language module*

The NLP module is concerned with reconstructing the vasculature and AVMs from text. This module does this by mapping natural language text into a meaning representation which can be used to extract a labelled graph from the knowledge base. We propose a computational model for translating medical reports into a meaning representation. The model is similar in spirit to that incorporated in the OSCON (Operating System CONsultant) system that answers English questions about computer operating systems (see Mc Kevitt 1986, 1991a, 1991b, 1992; and Mc Kevitt and Wilks 1987; and Mc Kevitt *et al.* 1992a, 1992b, 1992c, 1992d).

The NLP module will generate hypotheses about the component of vasculature under scrutiny independent of the VP module. Central to the NLP module is the meaning representation that is a formal description of the medical reports. The format of the meaning representation reflects the format of the medical report representing all of the immediate features mentioned.

We note that there are a number of standard phrases that the NLP module can search for in a medical report. First, there are standard phrases referring to specific objects such as AVMs or particular vessels. Second, there are references to attributes of these objects that describe their shape or size. Third, there are typically specifications of the spatial or topological relationships between objects. Finally, there are references to whether treatment is recommended or not. We propose the meaning representation template shown below. There is one root object:

[[Object(1) Attributes]
    [Object(2) Attributes]
    [Spatial/TopologicalRelationship]
    .
    .
    [Object(N) Attributes]
    [Spatial/TopologicalRelationship]
    [treatment]]

and many sibling objects:

[[Object(I) Attributes]
    [Object(J) Attributes]
    [Spatial/TopologicalRelationship]
    .
    .
    [Object(M) Attributes]
    [Spatial/TopologicalRelationship]]

This meaning representation is intended to be recursive in nature. We have constructed a set of rules that will map the medical text into this meaning representation.

The terms in the meaning representation will be matched to the KR in order to build a graph reconstruction from the medical report. This matching process will require a search that is analogous to that described for the vision module. The principle difference is the way in which individual matches are performed. In the vision module matches between vasculature are determined on a geometrical basis but here such matches are determined by comparison of text. We note also that the NLP module can facilitate learning of medical terminology.

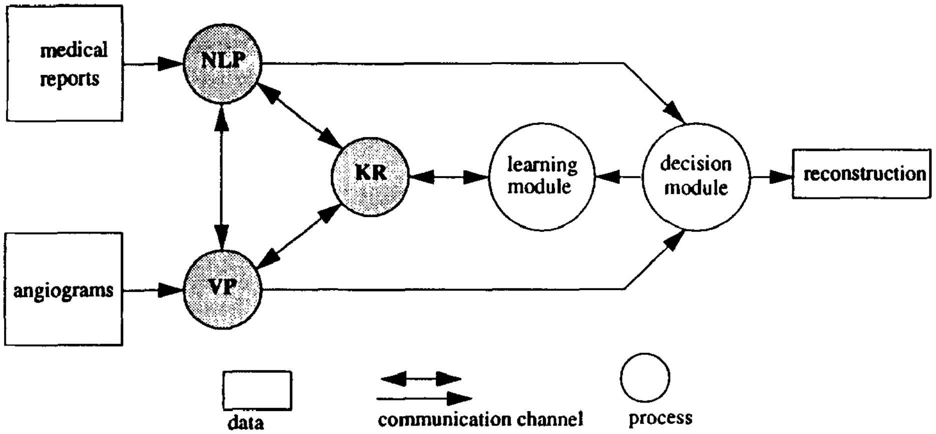


Fig. 4. The integrated system architecture.

3.4. *An integrating architecture*

The architecture of the integrated vision and language system is shown in Figure 4.

So far the KR, VP, and NLP modules have all been outlined. We now briefly discuss the operation of the decision module and the purpose of the communication channels.

Suppose the VP and NLP modules each worked independently to produce a set of reconstructions. Each of these reconstructions should be regarded as an hypothesis. We can discover where the modules agree by intersection of hypotheses. More generally we can consider Table III below.

Table III. Decision table for integrated language and vision.

$(G_{lang} \cap G_{vis})$	(G_{lang} / G_{vis})	(G_{vis} / G_{lang})	Row interpretation
∅	∅	∅	no reconstruction
∅	∅	match	vision only
∅	match	∅	language only
∅	match	match	no agreement
match	∅	∅	full agreement
match	∅	match	vision dominant
match	match	∅	language dominant
match	match	match	inconclusive

where G_{vis} and G_{lang} are the hypotheses from the VP and NLP modules respectively, ∅ is the empty set, and "match" denotes a non-empty set. Thus we see that the benefit of combining NLP and VP is that they disambiguate one another. Of the possibilities shown above the most useful has the interpretation "full agreement". In this case and when $(G_{lang} \cap G_{vis})$ is a singleton set we have an optimal result because then there is a unique solution. We argue that if the NLP and VP are allowed to communicate directly then this analysis will not be altered. This is because we have added no new information into the overall system so that no further ambiguity resolution is possible. Rather, the system will become more efficient.

4. CONCLUSION

We have described an architecture under which the power of NLP and VP are harnessed, via a single KR, for the purpose on reconstructing vasculature. Our contributions to this application are (1) a KR that is flexible and robust, (2) a novel reconstruction algorithm that is intended to yield more than one plausible result and (3) the proposal that medical reports can be used to assist reconstruction.

Future work involves closely addressing the issues of geometric matching and of learning geometrical shapes, implementing the natural language processor, and interfacing it to the visual processor.

Integrated models of vision and language processing will increase the speed and effectiveness of practical applications. More importantly, each processing module will help resolve ambiguities raised in the other. It is concluded that such investigations will not only aid in the development of practical systems but may shed light on the way in which humans integrate language and vision.

ACKNOWLEDGEMENTS

We would like to acknowledge: Ulf Bergvall of the Department of Radiology at the Royal Hallamshire Hospital; Lee Walton of the Department of Medical Physics at the Weston Park Hospital; Jim McGregor at the Department of Computer Science at the University of Sheffield; and Peter Andreae at Victoria University of Wellington for their comments on this work.

REFERENCES

André, E., Herzog, G. & Rist, T. (1988). On the Simultaneous Interpretation of Real-world Image Sequences and their Natural Language Descriptions: The System SOCCER. In Proceedings of *The 8th European Conference on Artificial Intelligence*, 449–454, Munich, Germany.

Ballard, D. H. & Brown, C. M. (1982). *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, U.S.A.

Baud, R. H., Rassinoux, A.-M. & Scherrer, J.-R. (1991). Knowledge Representation of Discharge Summaries. In Proceedings of *The Third Conference on Artificial Intelligence in Medicine (AIME 91)*, 173–182, Maastricht, June.

Delaere, D., Smets, C., Suetens, P. & Marchal, G. (1990). A Knowledge-based System for the Angiographic Reconstruction of Blood Vessels from Two Angiographic Projections. In Proceedings of *The North Sea Conference on Biomedical Engineering*, Nov. 1990.

Garreau, M., Coatrieux, J. L., Collerec, R. & Chardenon, C. (1991). A Knowledge Based Approach for Reconstruction and Labelling of Vascular Networks from Biplane Angiographic Projections. *IEEE Transactions on Medical Imaging* **10**(2): 122–131.

Hall, P. M. (1993a). *Problems in Reconstructing Vasculature and Lesions from Angiograms*. Technical Report. CS-93-05, Department of Computer Science, University of Sheffield, GB-S1 4DP, Sheffield, England, U.K., EU.

Hall, P. M. (1993b). *A Review of Procedures for Reconstruction of the Vasculature from Angiograms*. Technical Report CS-93-06, Department of Computer Science, University of Sheffield, GB-S1 4DP, Sheffield, England, U.K., EU.

Hall, P. M. (1994a). *Simulating and Animating Flow Through a Network of Tubes*. Technical Report CS-TR-94-13. Department of Computer Science, Victoria University of Wellington, Wellington, New Zealand. Or, via ftp.comp.vuw.ac.nz/doc/vuw-publications/CS-TR-94-13, or http://www.comp.vuw.ac.nz/Publications/CS-TR-94-13.

Hall, P. M. (1994b). *Simulating Angiography*. Technical Report CS-TR-94-15. Department of Computer Science, Victoria University of Wellington, Wellington, New Zealand. Or, via ftp.comp.vuw.ac.nz/doc/vuw-publications/CS-TR-94-15, or http://www.comp.vuw.ac.nz/Publications/CS-TR-94-15.

Hall, P. M. (1994c). *Simulating Angiography, a User Manual*. Technical Report CS-TR-94-17. Department of Computer Science, Victoria University of Wellington, Wellington, New Zealand.

Or, via ftp.comp.vuw.ac.nz/doc/vuw-publications/CS-TR-94-17, or http://www.comp.vuw.ac.nz/Publications/CS-TR-94-17. Source code available from: ftp.comp.vuw.ac.nz/pub/graphics/peter/xras-1.0.tar.gz.

Hall, P. M., Brady, J., Watt, A. H., Walton, L. & Bergvall, U. (1993). Segmenting and Reconstruction of Vascular Lesions from Biplane Angiograms. In Proceedngs of *DICTA 1993*, Sydney, Macquarie University, 802–809.

Hall, P. M. & McGregor, J. J. (1993). A Graph Based Model of a Collection of Physical Vasculature. In Proceedngs of *DICTA '93, Sydney, Macquarie University*, 414–421.

Hall, P. M., Andreae, P. & Ngan, M. (1995a). Reconstruction of Blood Vessel Networks from a Few Perspective Projections. *Artificial Neural Networks and Expert Systems '95*, Dunedin, New Zealand.

Hall, P. M., Feltham, R. & Fitzjohn, T. (1995b). Automated analysis of x-ray angiograms. *New Zealand Computer Society Conference*, 1995, Wellington, New Zealand. Also in Technical Report CS-TR-94-14, Department of Computer Science, Victoria University of Wellington, Wellington, New Zealand. Or, via ftp.comp.vuw.ac.nz/doc/vuw-publications/CS-TR-94-14, or http://www.comp.vuw.ac.nz/Publications/CS-TR-94-14.

Herzog, G., Sung, C.-K., André, E., Enkelmann, W., Nagel, H.-H. & Wahlster, W. (1989). Incremental Natural Language Description of Dynamic Imagery. In Freksa, C. & Brauer, W. (eds.), *Wissenbasierte Systeme. 3. Internationaler GI-Kongress*, 153–162. Springer-Verlag: Berlin.

Herzog, G. & Retz-Schmidt, G. (1990). Das System SOCCER: Simultane Interpretation und naturalichsprachliche Beschreibung zeitveranderlicher Szenen. In Perl, J. (ed.), *Sport und Informatik*, 95–119. Hofmann: Schorndorf.

Kingsland, L. C. (ed.) (1989). The 13th Annual Symposium on Computer Applications in Medical Care, *IEEE Computer Society Press*, November.

Kosslyn, S. M. & Pomerantz J. R. (1977). Imagery, Propositions and the Form of Internal Representations. *Cognitive Psychology* **9**: 52–76.

Kunii, T. L. & Shinagawa Y. (1991). New Concepts and techniques to Integrate Diverse Application Areas. In Patrikalakis, N. M. (ed.), *Scientific Visualization of Physical Phenomena, Proc. Computer Graphics International*, 3–24. Springer-Verlag: Tokyo.

Marr, D. (1982). *Vision*. Freeman: San Francisco.

Mayhew, J. E. W. & Frisby, J. P. (1981). Physcophysical and Computational Studies Towards a Theory of Human Stereopsis. *Artificial Intelligence* **17**: 349–385.

Maaß, W., Wizinski, P. & Herzog, G. (1993). *VITRA GUIDE: Multimodal Route Descriptions for Computer Assisted Vehicle Navigation*. Bereich Nr. 93, Universitat des Saarlandes, FB 14 Informatik IV, Im Stadtwald 15, D-6600, Saarbrücken 11, Germany, EU, February.

Mc Kevitt, P. (1986). *Building Embedded Representations of Queries about UNIX*. Memorandum In Computer and Cognitive Science, MCCS-86-72, Computing Research Laboratory, Dept. 3CRL, Box 30001, New Mexico State University, Las Cruces, NM 88003-0001, U.S.A.

Mc Kevitt, P. (1990). Acquiring User Models for Natural Language Dialogue Systems through Wizard-of-Oz Techniques. In Proceedings of *The Second International Workshop on User Modeling*, University of Hawaii at Manoa, Honolulu, Hawaii, U.S.A., March.

Mc Kevitt, P. (1991a). Principles and Practice in an Operating System Consultant. In Partridge, Derek (ed.), *Artificial Intelligence and Software Engineering, Vol. 1 Chapter on 'AI Mechanisms and Techniques in Practical Software'*, Ablex Publishing Corporation: New York.

Mc Kevitt, P. (1991b). *Analysing Coherence of Intention in Natural-language Dialogue*. Ph.D. Thesis, Department of Computer Science, University of Exeter, Exeter, GB- EX4 4PT, England, EU, September.

Mc Kevitt, P. (ed.) (1992). Natural Language Processing. *Artificial Intelligence Review* **6**(4): 327–332, Dec. Kluwer-Academic Publishers: Dordrecht, The Netherlands.

Mc Kevitt, P. & Wilks, Y. (1987). Transfer Semantics in an Operating System Consultant: The Formalization of Actions Involving Object Transfer. In Proceedings of *The Tenth International Joint Conference on Artificial Intelligence (IJCAI-87)* **1**: 569–575, Milan, Italy, EU, August.

Mc Kevitt, P., Partridge, D. & Wilks, Y. (1992a). *Analysing Coherence of Intention in Natural Language Dialogue*. Technical Report 227. Department of Computer Science, University of Exeter, GB- EX4 4PT, Exeter, England, EU.

Mc Kevitt, P., Partridge, D. & Wilks, Y. (1992b). *Why Machines should Analyse Intention in Natural Language Dialogue.* Technical Report 233. Department of Computer Science, University of Exeter, GB- EX4 4PT, Exeter, England, EU.

Mc Kevitt, P., Partridge, D. & Wilks, Y. (1992c). *Experimenting with Intention in Natural Language Dialogue.* Technical Report 234. Department of Computer Science, University of Exeter, GB-EX4 4PT, Exeter, England, EU.

Mc Kevitt, P., Partridge, D. & Wilks, Y. (1992d). Approaches to Natural Language Discourse Processing. *Artificial Intelligence Review* **6**(4): 333–364, Dec. Kluwer-Academic Publishers: Dordrecht, The Netherlands.

Mery, C., Normier, B. & Ogonowski, A. (1987). INTERMED: A Medical Language Interface. In Proceedings of *The Third Conference on Artificial Intelligence in Medicine (AIME 87)*, 3–8. Springer-Verlag: Berlin.

Neumann, B. & Novak, H.-J. (1986). *NAOS: Ein System zur naturalichsprachlichen Beschreibung zeitveranderlicher Szenen.* Informatik. Forschung und Entwicklung **1**(1): 83–92.

Partridge, D. (1991). *A New Guide to Artificial Intelligence.* Ablex Publishing Corporation: Norwood, New Jersey.

Pylyshyn, Z. (1993). What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery. *Psychological Bulletin* **80**: 1–24.

Rake, S. T. & Smith, L. D. R. (1987). The interpretation of x-ray Angiograms using a Blackboard Control Architecture. In Proceedings of *Computer Assisted Radiology 1987*, 681–685.

Ranum, D. L. (1988). Knowledge Based Understanding of Radiology Text. In Greenes, R. A. (ed.), *The 12th Annual Symposium on Computer Applications in Medical Care*, IEEE Computer Society Press, November.

Read, R. C. & Corneil, D. G. (1977). The Graph Isomorphism Disease. *J. Graph Theory* **1**: 339–363.

Retz-Schmidt, G. (1991). Recognizing Intentions, Interactions and Causes of Plan Failures. *User Modelling and User-Adapted Interaction* **1**: 173–202.

Retz-Schmidt, G. & Tetzlaff, M. (1991). *Methods for the Intentional Description of Image Sequences.* Bereich Nr. 80, Universitat des Saarlandes, FB 14 Informatik IV, Im Stadtwald 15, D-6600, Saarbrucken 11, Germany, EU, August.

Sager, N., Friedman, C. & Lyman, M. S. (1987). *Medical Language Processing: Computer Management of Narrative Data.* Addison-Wesley: Reading, MA.

Salamon, G. & Huang, Y. P. (1976). *A Radiological Anatomy of the Brain.* Springer-Verlag: Berlin.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures.* Lawrence Erlbaum Associates: Hillsdale, NJ.

Scherrer, J. R., Cote, R. A. & Mandil, S. D. (1989). (eds.) *Computerized Natural Language Processing for Knowledge Representation.* North-Holland: Amsterdam.

Schröder, M. (1992a). Knowledge-based Processing of Medical Language: A Language Engineering Approach. In Ohlbach, Hans-Jurgen (ed.), Preprints of the Proceedings of *The German Conference on Artificial Intelligence (GWAI-92)*, 190–199, Technical Report MP-1-92-232, Max-Planck-Institut fur Informatik, July.

Schröder, M. (1992b). Supporting Speech Processing by Expectations: A Conceptual Model of Radiological Reports to Guide the Selection of Word Hypotheses. In Gunther Gorz (ed.), *KONVENS 92-1 Konferenz, Verarbeitung naturlicher Spracher*, 119–128, Informatik aktuell, Nurnberg, October. Springer-Verlag: Berlin.

Schröder, M. (1992c). Knowledge based Analysis of Radiology Reports using Conceptual Graphs. In Pfeiffer, Heather (ed.), Proceedings of *The Seventh Annual Workshop on Conceptual Graphs*, 213–222. New Mexico State University (NMSU), Las Cruces, New Mexico, U.S.A., July.

Sowa, J.F. (1984). *Conceptual Structures: Information Processing in Mind and Machine.* Wesley: Addison.

Stansfield, S.A. (1986): ANGY: A Rule Base Expert System for Automatic segmentation of Coronary Vessels from Digital Subtracted Angiograms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**(2): 188–199.

Suetens, P., van Cleynenbreugel, J., Fierens, F., Smets, C., Oosterlink, A. & Wilms, G. (1987). An Expert System for Blood Vessel Segmentation on Subtraction Angiogram. *SPIE 767 Medical Imaging*, 454–459.

Wahlster, W. (1988). *One Word Says More than a Thousand Pictures*. On the Automatic Verbalization of the Results of Image Sequence Analysis. Bereich Nr. 25, Universitat des Saarlandes, FB 14 Informatik IV, Im Stadtwald 15, D-6600, Saarbrucken 11, Germany, EU, February.

Wilensky, R. (1983). *Planning and Understanding*. Addison-Wesley: Reading, MA.