A two layer case-based reasoning architecture for medical image understanding

Morten Grimnes and Agnar Aamodt

Department of Informatics
Norwegian University of Science and Technology
N-7034 Trondheim, Norway

Abstract. The paper describes a novel architecture for image understanding. It is based on acquisition of radiologist knowledge, and combines low-level structure analysis with high-level interpretation of image content, within a task-oriented model. A case based reasoner working on a segment case-base contains the individual image segments. These cases with labels are considered indexes for another case based reasoner working on an organ interpretation case base. Both are Creek type case based reasoners, here operating within a propose-critique-modify task structure. Methods for criticizing suggested interpretations by way of explanation, and how interpretations may be modified, are presented. An example run illustrates the system architecture and its key concepts.

1 Introduction

Image understanding has turned out to be a very difficult application task for AI methods. Methods exist that are able to do edge detection, and to some extent object identification, but methods for interpreting and understanding the content and meaning of whole pictures are less developed [Chellappa 92]. In the research reported here we address the interpretation of medical images, and in particular CT images. Computer Tomography (CT) is a widespread medical imaging technique which scans and displays a cross-section of the body using a type of X-rays. The images are mainly used for diagnostic purposes, as part of the overall patient examination procedure. The interpretation is therefore highly influenced by other patient data and the expectations they set up. We are exploring ways to combine lower-level segment identification with a higher level interpretation and understanding of the image as a whole. A segment, in this context, refers to an area of the image corresponding to an anatomical object or a significant part of one, for example an area picturing a liver.

At the core of our approach are two case-based reasoners, one for segment identification and one for the more wholistic interpretation of the image, corresponding to two layers of the system architecture. CBR covers a wide range of specific methods, ranging from syntactic pattern recognition type of methods [Wess 94][Aha 95], to knowledge-intensive methods oriented towards semantical and pragmatic contents [Leake 93][Aamodt 94a]. This makes CBR potentially suitable for both low and higher level image analysis, and for exploring the interfacing and the cooperative interpretation effort between the two layers. CBR methods at each of the two layers will accomplish different subtasks of the overall diagnostic task. This calls for a task oriented architecture, in which the medical diagnosis task is broken down into subtasks, and each subtask assigned to a particular part of the system. Methods are in turn assigned to each subtask, and each method specify the type and form of knowledge (cases as well as more general knowledge) that it requires. This results in an explicit problem solving and learning architecture targeted at the medical, diagnostic, image interpretation task.

Although simple diagnostic tasks best can be viewed as classification problem solving, the type of diagnosis we talk about here cannot. Detailed medical diagnosis involves planning of the diagnostic process as well as hypotheses construction, testing, reformulation, etc., making it more like a synthesis task than a pure classification one. In our case, the lack of a predefined, fixed set of image interpretations, and the complex process of arriving at an interpretation, has lead us to look into methods for construction and design problem solving to describe the overall image interpretation process. This type of method decomposes the top-level diagnostic image interpretation task into subtasks for which more detailed case-based methods in turn can be specified.

The focus of the paper is how explanation-driven case based reasoning methods (the Creek approach [Aamodt 91][Aamodt 94a]) fruitfully can be employed in the task of understanding abdominal CT images, within an architecture of two distinct but cooperating case bases. The next section gives a brief review of methods relevant for medical image understanding in general. Section 3 introduces the two-layered architecture within the framework of a generic design task model, and relate the subtasks to the reasoning methods within a Creek CBR system. In section 4 the task model of our system, named *ImageCreek*, is detailed, and the corresponding reasoning methods are described through an example. Discussion and status of research close the paper.

2 Methods for Tomographic Image Understanding

2.1 Medical image interpretation systems

There are a large number of medical image interpretation systems in the literature. Conventional image interpreting architectures [Swett 93][Gonzales 92] have also been employed in interpreting abdominal CT images [Englmeier 93]. Unfortunately, these architectures are characterized by performance brittleness and a lack of learning capability. A sufficient degree of system robustness as well as an ability to continuously learn from problem solving experiences are important for open-textured, weak-theory domains such as ours.

Beyond the conventional approaches, some systems focus on architectural aspects, others on knowledge modeling methods. Recent examples of architectural variations are a blackboard system with a hypothesize-and-test reasoning cycle for the radiological domain [Davis 91], an architecture based on genetic programming in which learning plays an important role [Teller 95], and a flexible architecture where image interpretation is viewed as a planning task [Gong 95]. Examples of systems where knowledge modeling plays an important role are the ERNEST system that has been applied to the interpretation of scintigraphic images and MRIs [Kummert 93], two systems on cranial MRI interpretation that feature knowledge modeling [Vernazza 87][Menhardt], a blackboard architecture featuring four diagnostic strategies in the radiological domain [Rogers 95], and a belief network based diagnostic decision aid on MRI liver lesions [Tombropoulos 94]. Characteristic of all the latter systems is that they lack a learning component.

In case-based reasoning there are a few examples of systems that are concerned with medical images. In the ROENTGEN system [Berger] therapy planning based on tomographic images is the object of reasoning. Radiologic image retrieval based on image captions and content related queries is the focus in the MacRad system [Macura 95]. None of these systems attempt to interpret images. The authors know of only one case based system that interprets medical images. This is the SCINA system [Haddad 95], which is based on the commercially available ESTEEM CBR tool. The system features a rule base for case adaptation and a case is represented as a matrix of integers. It seems the system features a limited possibility of knowledge modeling and is likely to experience the above mentioned brittleness problems.

Several support tools for general image analysis exist. In the implemented version of ImageCreek we are use the publicly available Khoros (version 2.0, from Khoral Research Inc.) image processing environment, for segmenting the image from the CT scanner.

2.2 Acquisition and modeling for abdominal CT understanding

Together with a radiology expert, a knowledge level analysis of the domain was made, as described in [Grimnes-96a]. In interpreting abdominal CT images radiologists tend to consider each projected anatomical entity separately, while at the same time combining the diagnostic hypotheses and findings for each entity into a complete image interpretation (see also [Wegener 92]). Wholistic aspects of the image and considerations pertaining to how well findings and hypotheses for each entity fit in with each other are crucial as well [Grimnes 96b].

Medical diagnosis in general is a complex type of diagnosis, which is tightly linked to examination procedures, diagnostic hypothesis formulation, testing and modification of hypotheses, etc. Often a preliminary diagnostic interpretation is arrived at which then has to be confirmed or rejected through a non-risky treatment regime. A task analysis of the abdominal CT interpretation problem reveals a type of problem and a set of tasks that in their scope and complexity to a larger degree conforms with solution synthesis than simple solution classification into a limited set of predefined categories. Our knowledge analysis and modeling approach is to a large extent based on the Components of Expertise framework [Steels-90], and some experiments performed within the Creest workbench [Winnem 96] (an extension of the KREST knowledge modeling tool [Steels 93] that incorporates the CreekL knowledge representation language). Our work has also been strongly influenced by the Generic Tasks [Chandrasekaran 93] and CommonKADS [Wielinga 93] methodologies. However, unlike the latter methodologies we put a strong emphasis on reducing the dependency on top-level modeling in favor of bottom-up, iterative model development - both in the system development phase and through sustained learning during system operation [Aamodt-95]. Our “world model” is decomposed into:

- A **Task** model where we principally look at the question of “what” is to be achieved and how specific tasks may be broken down into subtasks.
- A **Method** model where we look at how the tasks may be realized and how control flow is handled. In modeling methods we address the question of “how”.

- A **Domain Knowledge** model where we specify what kind of knowledge is necessary for achieving the tasks in terms of the available methods.

2.3 Integrated problem solving and sustained learning

The ImageCreek architecture is designed and implemented within the Creek system for knowledge-intensive case-based problem solving and learning. Cases, as well as general domain knowledge and information are captured in the frame-based representation language CreekL. A knowledge model represented in CreekL is viewed as a dense semantic network, where each node (concept) and each link (relation) in the network is explicitly defined in its own frame. A concept may be a general concept, a case, or a heuristic rule, and may describe domain objects as well as problem solving methods and strategies. The case-based method of Creek relies heavily on an extensive body of general domain knowledge in its problem understanding, similarity assessment, case adaptation, and learning.

The underlying case-based interpreter in Creek contains a three-step process of 1) activating relevant parts of the semantic network, 2) explaining derived consequences and new information within the activated knowledge structure, and 3) focusing towards a conclusion that conforms with the task goal. This “activate-explain-focus” cycle, is a general mechanism that has been specialized for each of the four major reasoning tasks of the CBR cycle [Aamodt 94b]. This is illustrated in Figure 1.

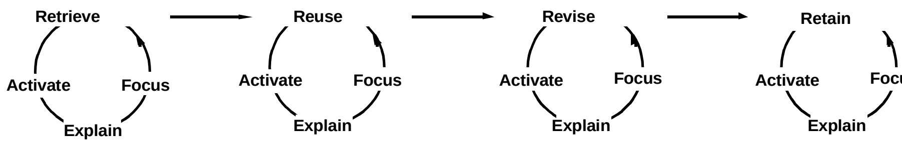

FIGURE 1. The CBR process and the explanation engine

The extensive, explanation-driven way of utilizing general domain knowledge in the CBR subtasks is a feature that distinguishes Creek from most other CBR systems.

A Creek system has the potential to learn from every problem solving experience. If a successful solution was directly copied from, or just slightly modified on the basis of a previous case, the reminding to that case from each relevant feature is strengthened, and no new case is stored. If a solution was derived by significantly modifying a previous solution, a new case is stored and difference links between the two cases are established. A new case is also created after a problem has been solved from rules or from the deeper knowledge model alone. The user is assumed to actively take part in both the problem solving and learning processes, e.g. by assessing hypotheses that the system cannot confirm or reject itself, supplying missing information, etc.

3 A system architecture for medical image understanding

3.1 A generic task model for medical image understanding

Inspired by work on Generic Design Tasks [Chandrasekaran 90], and in view of our knowledge acquisition findings, we have adapted a “*propose-critique-modify*” generic problem solving method. The method breaks down the top-level task of suggesting a diagnosis based on the CT image, into the four main subtasks:

- **Propose**. *Task* is to propose possible solutions. *Input* is a problem description and context. *Output* is a ranked list of possible solutions with normalized justifications. The solutions mark the spectre of likely and less likely solutions.
- **Verify**. *Task* is to verify that the solution(s) proposed fits the problem description. *Input* is a set of top ranked solutions and problem description. *Output* is a verified solution or failure with normalized justifications.
- **Critique**. *Task* is to look at what failed and propose a strategy for how the solution may be modified so as to better fit the problem description. *Input* is a normalized justifications on why verification failed, the failed solution in question and problem description. *Output* is a solution amendment strategy.
- **Modify**. *Task* is to realize the amendment strategy. *Input* is an amendment strategy, the failed solution, normalized justifications and problem description. *Output* is an amended solution.

3.2 A two layer approach to medical image understanding

In light of how radiologists seem to work and how current image processing algorithms are designed, we propose a two-layered architecture, corresponding to two case bases storing two different kinds of experience and supporting two different kinds of solutions. One layer lays on top of the other and both employ the *propose-verify-critique-modify* framework distinctly:

- The **Segment ImageCreek** Task layer (**SICT**). At this level we work with image segments (i.e. subsets of the image sharing some similarity) in isolation.

  i. In the ***case base***, a case is a description of a segment possibly together with a tentative pathologic/anatomical hypothesis as well as any previously rejected hypotheses with justifications. In this paper we use the term *segment hypothesis* denoting such a label (i.e. a case description of part of an image). Some of the hypotheses may be pathological labels, others may be normal anatomical labels and some may neither be pathological nor anatomical due to imaging or therapeutic artifacts.

  ii. The type of ***method*** associated with this layer is to only look at one segment description at a time and suggest a segment hypothesis of each single segment.

- The **Wholistic ImageCreek** Task layer (**WICT**). At this level we work with the entire image in question. We try to reach an overall image interpretation where all the different segments with suitable segment hypotheses fit in with each other and with the general problem description context.

  i. In the ***case base***, a case is a set of segments with diagnostic segment hypotheses together with the problem description not pertaining to single segments only, as well as any previously rejected segment hypotheses with justifications. In this paper we use the term *image interpretation* denoting such a case description.

ii. The type of ***method*** associated with this layer is to look at the broader aspects and the totality of the all the segment hypotheses in light of the problem description context and the findings not pertaining to a particular segment only.

4 The abdominal CT image understanding task

4.1 The generic model instantiated

We have instantiated the generic *“propose-verify-modify”* method with its *propose-verify-critique-modify* subtasks in light of the two case based reasoning layers described in section 3.2. The resulting task structure is illustrated in Figure 2. At the WICT level the object that is proposed, verified, possibly criticized and modified is an image interpretation. At the SICT level the object is a segment hypothesis. The input-output requirements are described in Table 1 based on the generic design in section 3.1.

There are two subtasks of the top ImageCreekTask (Figure 2) that will not be further described here. They are:

- **AcquireProblemDescription** - This task achieves three principal functions. It acquires all kinds of non-image findings like patient details and examination specifics. It furthermore plans an appropriate image segmentation procedure based on the non-image findings and finally executes the plan (segments the image). This task is outside the scope of this article.
- **ImageCreekLearn** - Top level task for the learning process in ImageCreek, corresponding to the Retrieve task in Figure 1. This task is not covered in this article, but briefly it is a combination of failure driven learning and case integration initiated at the end of a case based reasoning cycle. The expert may be requested to volunteer new knowledge in case of reasoning problems.

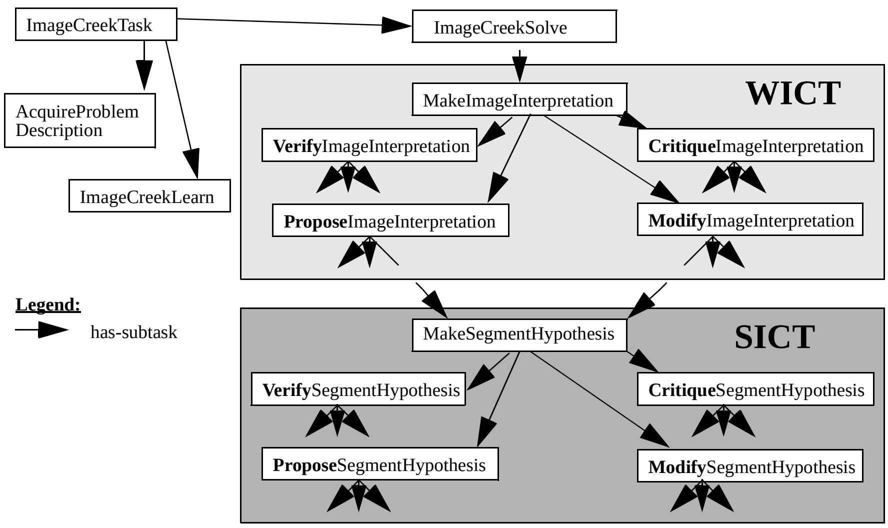

FIGURE 2. The two layer propose-verify-critique-modify task hierarchy in ImageCreek

Task	Input	Output
ProposeImageInterpretation	Input image interpretation case findings, Domain Knowledge	Ranked list of previous image interpretations (cases), context activated knowledge, normalized justifications
ProposeSegmentHypothesis	Input segment case findings, Domain Knowledge	Ranked list of previous segments with hypotheses (cases), context activated knowledge, normalized justifications
VerifyIImagenterpretation	Proposed image interpretation cases, input image interpretation case, context activated knowledge	Normalized justifications and verification decision
VerifySegmentHypothesis	Proposed segmentation cases, input segmentation case, context activated knowledge	Normalized justifications an verification decision
CritiqueImageInterpretation	Normalized criticism, proposed image interpretation case, input image interpretation case	Modification strategy or failure
CritiqueSegmentHypothesis	Normalized criticism, proposed segmentation case, input segmentation case	Modification strategy or failure
ModifyImageInterpretation	Modification strategy, ranked list of proposed image interpretations, input image interpretation case, normalized justifications	New interpretation case or failure with resegmentation flag set or unset
ModifySegmentHypothesis	Modification strategy, ranked list of proposed segments with segment hypotheses, input segment case, context activated knowledge	New segment hypothesis or failure

TABLE 1. Input and output for the Propose, Verify, Critique and Modify tasks

4.2 The SICT-WICT interface

The interaction between the segment and wholistic level is a core issue in the architecture (see Figure 3). The interaction takes place basically through two interfaces:

1. In **proposing an image interpretation**: We achieve ProposeImageInterpretation by first generating possible image interpretations (GenerateInterpretations). Generation is achieved by direct reminding to earlier cases (ProposeInterpretationFromInterpretationCases) and by hypothesizing new diagnostic labels for all the segments output from the image processing subsystem (ProposeInterpretationBySegmentIntegration). The process of finding diagnostic segment hypotheses for the

various segments is the main task of SICT (section 3.2). The MakeSegmentHypothesis task is the main SIC Task whereas MakeSegmentHypotheses is a task that combines a set of segments that are assigned the same diagnostic segment hypothesis into one complex entity.

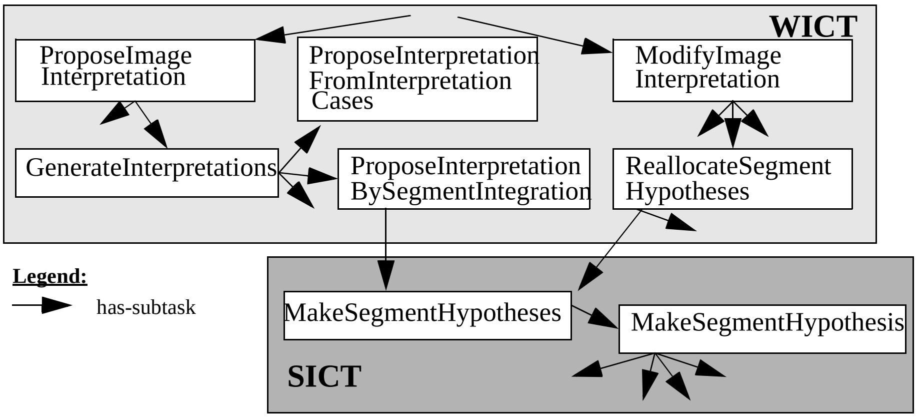

FIGURE 3. The SICT-WICT interface

2. In **modifying an image interpretation**: One way of modifying a solution is to alter parts of it. This is the task of ReallocateSegmentHypothesis. ReallocateSegmentHypotheses looks at the justifications of the image interpretation and bases it´s judgement on how well a segment hypothesis seems to fit in on this. The ones having the weakest justifications are picked, and MakeSegmentHypotheses is requested to suggest new segment hypotheses for these.

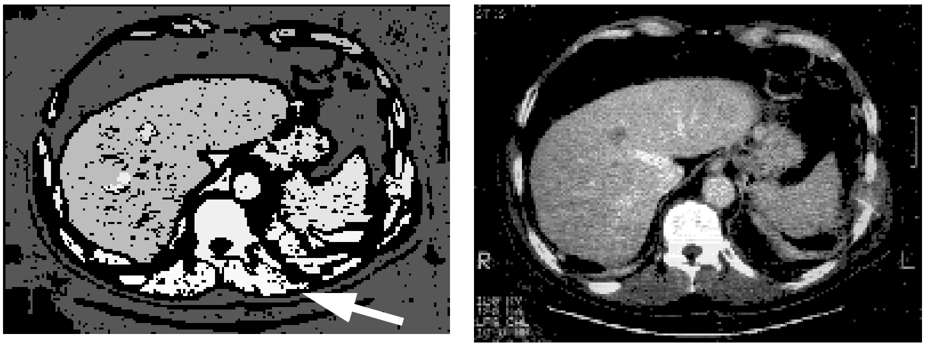

FIGURE 4. The segmented version (left) of an example input image (right). Arrow indicates example segment (section 4.3)

4.3 Proposing an image interpretation.

Below we illustrate key aspects of this reasoning process by presenting excerpts from an example run of ImageCreekSolve. Let´s assume we have acquired a problem description (AcquireProblemDescription in figure 2) of the right image in Figure 4. The description is the input problem case to the ImageCreek system and includes general non-image findings pertaining to the patient and examination:. The image segmentation produced by Khoros comes as additional input.

icase-b1-1-t001

...		...	
has-sex	male	has-exam-modality	ct
has-age	seventies-years	has-exam-machine-id	rit-old-ct
has-earlier-diagnosis	lung-cancer	has-contrast-injection-type	prolonged
has-earlier-diagnosis-time-ago	ayear-threeyears-ago	has-contrast-time	portovenuous
has-social-condition	retirement-home	has-ct-slicethickness	ten-mm
has-habit	smoker	has-ct-angle	zero-degrees
...		...	

Image interpretations are suggested by the subtask GenerateInterpretations (Figure 3). The task is achieved in two principal ways, through case remindings from findings - so that ProposeInterpretationFromInterpretationCases is able to retrieve a set of earlier cases, or by hypothesizing a diagnostic label for each image segment. Let us take a closer look at the latter:

Based on the problem description, a number of segments cases are retrieved. One of the cases looks like:

scase-b1-1-42-bp

instance-of	no-heart-or-kidney-level-scase	has-ventral-bb	dorsal
has-size	spine-size	has-intensity	sixties-hu
is-left-right	narrow-lr-centre	has-stdev	stdev-large-hu
is-ventral-dorsal	dorsal	has-contrast-injection-type	prolongued
has-orientation	left-right	has-contrast-time	portovenuous
has-eccentricity	slightly-eccentric	has-exam-purpose	routine-search-for-abnormality
has-left-bb	left-centre	has-ct-bodysection	nokidney-noheart-level-region
has-right-bb	narrow-lr-centre		

ImageCreek does two things to each of these segment cases early in the ProposeImageInterpretation process:

1. Each segment receives a diagnostic hypothesis. This is the overall goal of SICT and in our example the segment scase-b1-1-42-bp is assigned this hypothesis (MakeSegmentHypothesis achieves this):

   has-segment-hypothesis c-nok-noh-left-erector-spinae

2. In each segment there are certain findings that are relevant to the interpretation of the entire image. These are extracted and integrated into the input image interpretation case. This is achieved by ProposeInterpretationBySegmentIntegration. The *transformed* **scase-b1-1-42-bp** then looks like:

...		...	
has-left-erector-spinae-scase	(scase-b1-1-42-bp)	has-left-erector-spinae-left-bb	left-centre
has-left-erector-spinae-hypothesis	c-nok-noh-left-erector-spinae[a]	has-left-erector-spinae-right-bb	narrow-lr-centre
is-left-erector-spinae-left-right	narrow-lr-centre	has-left-erector-spinae-ventral-bb	dorsal
is-left-erector-spinae-ventral-dorsal	dorsal	has-left-erector-spinae-dorsal-bb	extreme-dorsal
...		...	

a. The erector spinae is a muscle adjacent to the spine. The abbreviation c-nok-noh is simply an abbreviation of contrast-no-kidney-no-heart-level which is indicates the entity´s position in an anatomical classification hierarchy.

Findings pertaining to the image segments undergo several transformations, some in AcquireProblemDescription, others in MakeSegmentHypotheses. They therefore-rely on image segmentation functions. Further, they get associated with a medically relevant segment hypothesis. This has consequences:

- We cannot simply accept the findings as true by default. Findings pertaining to image segments must be **questioned** and possibly **rejected**
- In rejecting findings we need a mechanism for **evaluating the quality of the findings** as well as **handling the rejection**. This is partly what we try to do in CritiqueImageInterpretation and ModifyImageInterpretation respectively.

4.4 Testing and Verifying an image interpretation

In ProposeImageInterpretation the generated interpretations are tested for relevance in terms of explaining how serious syntactic differences between solution (retrieved cases) and problem description (input case) are. In VerifyImageInterpretation we generate expected findings of the retrieved case and explain how these are met in the input case. This task, as well as the other high-level tasks, have their own detailed subtask structure not elaborated here.

The explanation process is handled by a particular task (the GenericExplanationTask) which uses the domain knowledge base and a context-based spreading mechanism (a type of bounded beam search) and search from a set of support concepts to a set of concepts to be explained. The search results in an explanation path with a certain combined explanation strength. This is a computed value, where the most important role is played by the explanatory strength of each relation that is part of the path.

- An example: In the transformed scase (above) a represented segment was hypothesized to be a left-erector-spinae which is not present in the retrieved scase. In light of the retrieved findings, can we justify that left-erector-spinae should be present as a segment hypothesis? A justification task (JustifyInputFindings), using the GenericExplanationTask, gives us:

(((START C-NOK-NOH-LEFT-ERECTOR-SPINAE) (IS-COMPLETELY-DORSAL-OF C-NOK-NOH-COLON) (IS-COMPLETELY-LEFT-OF C-NOK-NOH-RIGHT-ERECTOR-SPINAE)) 0.8640000000000001)

The explanation structure as shown is a dump from an execution run. It shows an explanation chain, built up of concept-relation-concept triplets, and assigned the resulting explanation strength as a numerical value.

All explanations are stored as part of the image interpretation case. A decision is made whether to accept the image interpretation or reject it. In our example it is accepted and will be subject to verification.

A part of the verification task is to look at the solutions proposed (essentially the set of segment hypotheses generated by ProposeInterpretationBySegmentIntegration), generate expectations from these and justifying whether these are met.

- In our example one of the retrieved segment hypotheses were right-erector-spinae. The GenericExplanationTask states that given a right-erector-spinae, we should expect a spine as well:

  (((START C-NOK-NOH-RIGHT-ERECTOR-SPINAE) (IS-COMPLETELY-RIGHT-OF C-NOK-NOH-COLON) (IS-COMPLETELY-VENTRAL-OF C-NOK-NOH-SPINE)) 0.7635266498559999)

  This and other expectations are summed up in the list of expected findings:

  ((C-NOK-NOH-RIGHT-ERECTOR-SPINAE 0.8640000000000001) (C-NOK-NOH-FIRST-LEFT-RIB 0.8) (C-NOK-NOH-COLON 0.8) (C-NOK-NOH-SPINE 0.7635266498559999) (C-NOK-NOH-SPLEEN 0.8552960000000001) (C-NOK-NOH-AORTA 0.7672) (C-NOK-NOH-LEFT-ERECTOR-SPINAE 0.8) (C-NOK-NOH-GALL-BLADDER 0.88) (C-NOK-NOH-LIVER 0.7031799999999999))

  The task JustifyExpectedFindings receive the list of expected findings not occuring in the input case, and try to explain how they are similar to the input findings, in a similar manner to what JustifyInputFindings earlier did.

4.5 Criticizing an image interpretation

In the present version of ImageCreek this function exists only as an early design. The task is split into two subtasks:

- **DiagnoseInterpretationFailur**e: The method we chose looks at the source, the severity and kind of criticism in classifying the failure:

  i. If criticism is relatively even on all segments tag the criticism as **Even**. Otherwise, tag it **Uneven**.

  ii. If the total of criticism is very strong, tag the criticism **Strong**. Otherwise tag it **Moderate**.

  iii. If the criticism is strong from the ProposeImageInterpretation step tag the criticism **StructuralCriticism** otherwise tag it **SolutionCriticism**.

- **Proposing an amendment strategy**: Some tentative rules we propose are:

  i. In case criticism is **Uneven** or **Moderate** propose segment reallocation.

ii. In case criticism is **Even** or **Strong** propose finding a new image interpretation.

iii. In case criticism is **StructuralCriticism** count this as an argument for a finding a new image interpretation.

iv. In case criticism is **SolutionCriticism** count this as an argument for a proposing segment reallocation.

v. If this is second (or more) time verification fails from segment reallocation, propose finding a new image interpretation.

vi. If this is second (or more) time verification fails from finding new image interpretations, propose image resegmentation.

4.6 Modifying an image interpretation

ModifyImageInterpretation have four ways of attempting to modify the failed and verified image interpretation.

1. From ModifyImageInterpretation we receive the collected criticism of the verified image interpretation. Each finding pertaining to a segment is related to a particular segment hypothesis (e.g. has-left-erector-spinae-dorsal-bb is related to the c-nok-noh-left-erector-spinae hypothesis. All findings pertaining to such a segment hypothesis are grouped together and a total criticism factor for that particular segment hypothesis is calculated.
2. If an identified segment hypothesis has a corresponding set of input segment cases these are identified and the segments together with the segment hypothesis is added to the CandidateSegments set. The segment hypothesis is included and named the ForbiddenHypothesis so that MakeSegmentHypotheses shall know that it must not associate the particular segments with this segment hypothesis again.
3. All identified segment hypotheses are named CandidateHypotheses. Control is left to MakeSegmentHypotheses.

As control is resumed the new segment hypotheses are integrated into the input image interpretation case and control is left to ModifyImageInterpretation which will leave control to ImageCreekSolve, which in turn will make the input case subject for renewed verification.

5 Discussion

The ImageCreek architecture crucially depends on a number of assumptions to ensure robustness and problems solving quality. Among the more important are:

1. **Case representation**. The findings in segment cases must show to be well chosen in that they (i) ensure a reasonable feature cluster separability (linear separability is not necessary) and (ii) serve as relevant indices such that the Propose subtasks at WICT and SICT level are able to generate a high quality set of solutions.

2. **Image segmentation robustness**. The correspondence between an image segment and an anatomical entity must be reasonably true. That means that it must not happen too frequently that there is no such one-to-one correspondence. If this happens the system will gracefully degrade in performance. To ensure that this one-to-one correspondence is generally true, a method of intelligent planning of image segmentation processing must be devised and the underlying image processing functions must be reasonably stable.

3. **Learning capability**. Not only must revised cases be retained for future retrieval but there must be a method correcting problem solving anomalies and integrating new knowledge. The ImageCreek prototype knowledge base is already less than trivial to maintain and to be able to make the architecture scalable, methods for knowledge base maintenance are crucial.

The principal strong sides of the ImageCreek approach are:

- **Domain knowledge language.** ImageCreek relies heavily on explanations in problem solving. The explanations are based on the domain knowledge. Care has been taken to make a knowledge base ontology that is as close to the radiologist´s language as possible. This ensures a quality control in that the radiologist more easily can verify the explanations from ImageCreek and that learning indices and from failures can be evaluated by a radiologist with minimum ImageCreek experience.
- **Context sensitive problem solving.** Our knowledge acquisition shows that the diagnostic image interpretations of a radiologist are highly sensitive to findings pertaining to the patient´s general characteristics, referring physician, clinical history and current diagnoses and medications. The ImageCreek system ensures that such concerns appropriately influences the image interpretation process.
- **Integrated problem solving and learning.** Every solved problem is an experience to learn from in ImageCreek. Since we base our architecture on Creek methods we profit from the methology´s ability to learn and solve in tandem. Problem solving and learning is involving the user interactively.

Looking at the literature we feel that the systems that are closest in having appropriate solutions for our domain are the ERNEST system [Kummert 93] and the VIA-RAD system [Rogers 95]. Both are realized and well evaluated. ERNEST seem to scale up reasonably well, is fairly close to the expert in terminology. The system has a way of explaining it´s actions but these justifications are mere user presentations and do not influence the problem solving itself. Control is handled in an elegant and generic fashion. In VIA-RAD much the same is the case. It is unclear as to how well it would scale up and it features a domain knowledge language that is fairly close to the expert´s. The systems´ principal weakness are their lack of learning from experience. It is unclear to what an extent either system would be sensitive to problem solving context.

6 Status

Currently the problem solving part of the architecture is past the early design stage and the key aspects are implemented in a prototype. The image processing subsystem is in a similar stage. ImageCreekLearn is in an earlier design stage where initial implementation is due to begin soon. The image processing subsystem runs on a Unix Solaris system and is an integration of c based Khoros library routines and custom made shell scripts. The other part of the prototype system is written in a combination of the CreekL representation language and standard Common Lisp for Macintosh.

Image understanding is a difficult task, where methods for integration of low-level and higher-level image analysis and interpretation is called for. This represents a challenge for both the AI and the image understanding communities. Given the problem of developing sufficiently strong general domain knowledge models in these domains, a case-based approach is a plausible suggestion. The results of our research so far clearly indicates that this also is a promising approach.

7 Acknowledgment

We would like to thank Pinar Øzturk for valuable input on aspects of the architectural model. Particularly the ideas on the Generic Explanation Task are inspired by her work. Furthermore we are indebted to Arnt Lockert, Ståle Heitmann and Ole Martin Winnem for implementing a number of Creek methods and the Creest modeling tool. Finally we acknowledge the invaluable input of our collaborators at the Regional Hospital of Trondheim, in particular Medical Drs. Asbjørn Ødegård and Klaus Schipper.

8 References

[Aamodt 91] Aamodt A., *A Knowledge-Intensive, Integrated Approach to Problem Solving and Sustained Learning*, Ph.D. thesis, May 1991, Dept. of Computer Systems and Telematics, University of Trondheim, Norway

[Aamodt 94a] Aamodt A., Explanation-Driven Case-Based Reasoning, pp. 274-288 in *Topics in case-based reasoning* by Wess S. et al. (eds.), Springer 1994, pp. 274-288.

[Aamodt 94b] Aamodt A. and Plaza E, Case-Based Reasoning: Foundational issues, methodological variations, and system approaches, *AI Communications*, Vol 7, no.1, pp. 39-59.

[Aamodt 95] Aamodt A., Knowledge acquisition and learning by experience - the role of case specific knowledge, in *Machine Learning and Knowledge Acquisition*, Academic Press 1995, ISBN 0-12-685120-4, pp. 197-245.

[Aha 95] Aha, D., Weighting features, In Veloso, M. and Aamodt, A. (eds.) *Case-based reasoning research and development*, First International Conference, ICCBR-95, Sesimbra, Portugal, October 1995, Springer Verlag, 1995, pp 347-358.

[Berger] Berger J., Roentgen: Radiation Therapy and Case-based Reasoning, Artificial Intelligence Laboratory, University of Chicago, Chicago, Illinois 60637.

[Chandrasekaran 90] Chandrasekaran B., Design Problem Solving: A Task Analysis, *AI Magazine*, Winter 1990.

[Chandrasekaran 93] Chandrasekaran B. and Johnson T., Generic Tasks and Task Structures: History, Critique and New Directions in David, J.-M. et al. (eds.),*Second Generation Expert Systems*, Springer 1993, ISBN 3-540-56192-7.

[Chellappa 92] Chellappa R and Kashyap, R.L., Image understanding, In Shapiro, S.C (ed.), *Encyclopedia of artificial intelligence,* Vol 1, John Wiley & Sons, 1992pp 641-663.

[Davis 91] Davis D. and Taylor C., A Blackboard Architecture For Medical Image Interpretation, *SPIE* Vol. 1445 Image Processing (1991).

[Englmeier 93] Englmeier K. et al., Model Based Image Interpretation of Spiral CT Scans of the Abdomen, pp. 44-51, in *Artificial Intelligence in Medicine* by Andreassen S., Engelbrecht R and Wyatt J. (eds.), IOS Press 1993.
[Gonzales 92] Gonzales R. and Woods R., *Digital Image Processing*, Addison Wesley 1992, ISBN 0-201-50803-6.
[Grimnes 96a] Grimnes M. and Ødegård A., Knowledge acquisition for robust image understanding: Knowledge use in CT image understanding, Report at dept. of Informatics, NTNU, ftp://sophus.ifi.unit.no/YarcWorld/public/papers-reports/kacq.kn-use.19-01-96.ps
[Grimnes 96b] Grimnes, M., Knowledge breadth in CT image diagnosis, Submitted for publication, 1996.
[Gong 95] Gong L. and Kulikowski C., Composition of Image Analysis Processes Through Object-Centered Hierarchical Planning, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 10, October 1995.
[Haddad 95] Haddad M. et al., SCINA: A Case-Based Reasoning System for the interpretation of Myocardial Perfusion Scintigrams, Dept. of Cardiology, University of Vienna, AKH Wien, Vienna, Austria, email: haddad@vm.akh-wien.ac.at
[Kummert 93] Kummert F. et al., Control and explanation in a signal understanding environment, *Signal Processing* 32 (1993) 111-145
[Leake 93] Focusing construction and selection of abductive hypotheses. In *IJCAI*, *1993*, pp. 24-31. Morgan Kaufmann.
[Macura 95] Macura R. and Macura K., MacRad: Radiology Image Resource with a Case-Based Retrieval System, pp. 43-54, in Veloso M. and Aamodt A. (eds.), *Case-Based Reasoning Research and Development*, Springer 1995, ISBN3-540-60598-3
[Menhardt] Menhardt W. et al., Knowledge Based Interpretation of Cranial MR Images, Philips GmbH, Forschungslaboratorium Hamburg POBox 540840 D-2000 Hamburg 54, FRG
[Rogers 95]þRogers E., VIA-RAD: a blackboard-based system for diagnostic radiology, *Artificial Intelligence in Medicine* 7 (1995), pp. 343-360
[Schalkoff 92] Schalkoff R., *Pattern Recognition. Statistical, Structural and Neural Approaches*, Wiley 1992, ISBN 0-471-55238-0
[Steels 90] Steels L., Components of Expertise, *AI Magazine* summer 1990.
[Steels 93] Steels, L., The component framework and its role in reusability, In David, J-M. et al. (eds.), Springer 1993, ISBN 3-540-56192-7, pp. 273-298.
[Swett 93] Swett H. et al., Computer Vision and Decision Support, pp. 273-313 in *The Perception of Visual Information* by Hendee W. and Wells P. (eds.), Springer 1993
[Teller 95] Teller A. and Veloso M., PADO: A new learning architecture for object recognition, *Symbolic visual learning*. Oxford University Press, 1995.
[Tombropoulos 94] Tombropoulos R. et al., A Decision Aid for Diagnosis of Liver Lesions on MRI, Section on Medical Informatics, Stanford University School of Medicine, 1994 AMIA
[Vernazza 87] Vernazza G.L. et al., A Knowledge-Based System for Biomedical Image Processing and Recognition, *IEEE Transaon Circuits and Systems*, Vol. cas-34, no. 11, November 1987
[Wess 94] Wess, S., Althoff, K-D., Derwand, G., Using k-d trees to improve the retrieval step in case-based reasoning, In Wess. S., Althoff, K-D., and Richter, M.M. (eds.) *Topic in case-based reasoning*, Springer Verlag, 1994, pp. 167-181.
[Wegener 92] Wegener O., *Whole Body Computed Tomography*, 2nd ed., Blackwell scientific pub. 1992, ISBN 0-86542-223-0
[Wielinga 93] Wielinga B. et al., Towards a Unification of Knowledge modeling Approaches, in David, J-M. e al. (ed), *Second Generation Expert Systems*, Springer 1993, ISBN 3-540-56192-7.
[Winnem 96] Winnem O., Integrating knowledge level and symbol level modeling - The Creest Workbench, MSc thesis, Dept. of Informatics, Norwegian University of Science and Technology - NTNU, Trondheim, 1996.