From Description to Decision:

Towards a Decision Support Training System for MR Radiology of the Brain

Benedict du Boulay[1], Briony Teather[2], George du Boulay[3], Nathan Jeffrey[2], Derek Teather[2], Mike Sharples[4], and Lisa Cuthbert[1]

[1] School of Cognitive and Computing Sciences, University of Sussex
[2] Department of Medical Statistics, De Montfort University, Leicester
[3] Institute of Neurology, London
[4] School of Electronic and Electrical Engineering, University of Birmingham

Abstract. We have developed a system that aims to help trainees learn a systematic method of describing MR brain images by means of a structured image description language (IDL). The training system makes use of an archive of cases previously described by an expert neuroradiologist. The system utilises a visualisation method – an Overview Plot – which allows the trainee to access individual cases in the database as well as view the overall distribution of cases within a disease and the relative distribution of different diseases. This paper describes the evolution of the image description training system towards a decision support training system, based on the diagnostic notion of a "small world". The decision support training system will employ components from the image description training system, so as to provide a uniform interface for training and support.

1 Introduction

We have developed an image description training system [19, 20] that aims to help radiology trainees learn how to describe MR brain images in a systematic way by means of a structured image description language (IDL). This language allows clinically meaningful features of MR brain images to be recorded, such as the location, shape, margin and interior structure of lesions. The training system makes use of images from an archive of about 1200 cases, previously described in detail using the terms of the IDL by an expert neuroradiologist.

The image description training system employs a visualisation method – an Overview Plot – which allows the trainee to view and access (i) the images themselves, (ii) the written descriptions of the individual lesions in the image, and (iii) a two dimensional representation of the multi-dimensional distribution of all cases of a disease chosen from the archive. The two dimensional representation relates to, and is calculated from, the descriptions of the lesions. Thus one can view the overall distribution of appearance of cases within a disease and the relative distribution of different diseases, one against another. To this extent it is a kind of case-based training system that provides a visual indexing mechanism to cases similar to the case in hand.

This paper describes the development of the image *description* training system to include a second stage, namely a *decision support* training system, which we see as its immediate role with future potential as a decision support tool.

Two principles have guided the development of the system so far. First, the system is deliberately aimed to support and train the radiologist's inferences from what can be observed *in the images*. In particular, the two dimensional representation is currently based on lesion appearance and confirmed diagnosis, but not on clinical signs and symptoms. The reconciliation of those inferences with other sources of data, such as the clinical history, is a matter for the user. Second, the design exploits as far as possible radiologists' visual-spatial reasoning rather than simply offering numerical or quasi-numerical information about diagnostic probabilities.

In the next section we briefly outline the nature of radiological expertise as it informs the design. A comparison is then made with other knowledge-based learning and teaching environments for radiology that offer substantial adaptivity to the individual or are based on a careful analysis of the training task. The body of the paper briefly describes the image description language and the overview plot, and then outlines the decision support methodology. It concludes with a discussion of our initial evaluation of the component tools and future work.

2 Background

2.1 Medical and Radiological Expertise

Medical experts possess highly structured knowledge that informs the small set of hypotheses that need to be considered in order to make accurate diagnoses. Their reasoning is generally data driven [16] and does not appear to work directly from scientific first principles so much as from an "illness script" that encapsulates various levels of knowledge (including, at base, the scientific) in a schema associated with a particular pathology [18]. When presented with a new case experts rapidly home in on a number of "critical cues" that guide them to consider that small set of possible hypotheses which best explains the data (a "small world") [11, 12]. Experts are also strongly guided by "enabling conditions", i.e. crucial factors in the patient data or clinical history. Experts have schemata that are augmented with vivid, individual cases that they have seen and use these in dealing with new cases [8]. Experts have an excellent appreciation for the range of normality but have a propensity to pay attention to and recall abnormal cases better than normal ones [9].

Expert radiologists are able to identify much of the abnormality in an image very quickly (an initial gestalt view) and this is followed by a more deliberative perceptual analysis, though both stages incorporate data-driven and hypothesis-driven activity [2]. More importantly they have undergone a combined perceptual/conceptual change, evolving from recognising salient image intensities towards recognising diagnostically significant image features. They are better than novices at identifying the 3D location and physical extent of the abnormality (i.e. responding to "localisation cues") [13].

Both experts and novices are sensitive to the skewing effect on diagnosis of other information about the patient [15]. Consulting this prior to viewing the images affects not only what they see but also what they diagnose and therefore recommend. This raises the difficult issue of when in the analysis the radiologist should look at the clinical data and case history.

2.2 Computer-based Training

While there are many computer-based training aids for radiology (including neuroradiology), most are essentially electronic books or collections of images together with some kind of indexing mechanism, normally based primarily on disease. There have been relatively few systems that attempt to either model the domain or the evolution of knowledge and skill of the student in a detailed way. Of these, Azevedo and Lajoie [2] describe an analysis of the problem solving operators used in mammography as applied by radiologists of various levels of skill. They also analyse the nature of teaching as it occurs in radiology case conferences and particularly the way that experts articulate their diagnostic reasoning. Both these analyses are used as part of the design process for RadTutor [1]. A similar careful analysis in the domain of chest X-rays has been carried out by Rogers [17] as part of the design process of VIA-RAD tutor. Macura and Macura and their colleagues [14] have taken a case-based approach that is similar to our own in a tutor for CT and MR brain images. Their system offers a case-retrieval and decision-support mechanism based on descriptors but does not employ a detailed image description language nor offer an overview plot. However their system does employ an atlas and contains tutorial material and images of normal brains as well as those displaying lesions. It can act as a decision support system by offering a range of possible diagnoses and access to the images of related cases, given the textual information that has been entered.

3 Visual Decision Support Training

3.1 Image Description Language

The basic domain representation underpinning the system is an archive of cases with confirmed diagnoses, all described by the same expert (G. du Boulay) using the IDL. These include separate descriptions for each image sequence/echo as well as detailed descriptions (e.g. the *region, major position, exact location, margin, structure, shape, area, conformity to anatomical feature, interior pattern* (if any) and its *intensity*) of the lesion (or the largest of each type of lesion visible, where there are multiple lesions), as well as *correspondence* between described parts of lesions seen under different sequences and descriptions of *atrophy*, *other signs* and other *abnormal signals* for the case as a whole [4].

The image description language for MR was derived using an iterative prototyping approach, utilizing experience gained in a similar enterprise for CT brain images and a menu-based computer advisor (BRAINS) to aid in image interpretation and cerebral disease diagnosis [21, 22].

Subsequent to the process of validation and refinement, G. du Boulay employed an interactive description tool (MRID, running under X-Windows for Unix workstations) to describe an archive of some 1200 cases using the terminology of the IDL. These represent a sample of the abnormal cases captured at two different imaging centres dealing with very varied disease.

The IDL describes the appearance of the images rather than the underlying disease, though the ontology of the language is influenced by a knowledge of diagnostically important disease processes. The IDL has been constructed to be as complete and detailed as possible, taking account of the wide range of diagnostic problems that occur in neuroradiology and the variation of image appearance according to sequence type. It should be noted that one of the difficulties found in earlier work is still of major importance. The process of exhaustive description is long and painstaking, and the more recent gains in selecting terms by menu on a computer screen are offset by the more extensive and detailed descriptors made possible by MRI.

For the purposes of the prototype description training system a simplified version of the description language has been used. It provides an initial set of terms to support discussion and sharing of knowledge amongst trainee neuroradiologists and their supervisors. It also serves as a structured representation of knowledge for the MR Tutor, enabling it to generate remedial responses to student errors.

3.2 Display of Small Worlds

We can consider a case as occupying a point in a many-dimensioned space of description features. For the simplified language this space has some 30 dimensions, where each point is a vector of binary values, each representing the presence or absence of a particular feature[1]. Multiple Correspondence Analysis (MCA) is a statistical technique for data reduction and visualisation [7]. It is used here to reduce the dimensionality down to two so as to provide a ready means of overviewing the data. It does this by finding that plane which best spreads out the subset of cases under consideration. MCA is similar to principal components analysis but is applied to categorical/binary data as opposed to scalar data and assesses all possible pairwise associations in the data. Whilst the technique treats ordinal values such as as "tiny", "small", "medium" or "large" as separate dimensions, it has the advantage of not depending on the allocation of arbitrary scale values to these categories.

Effectively, a set of X-Y weightings for each feature value is derived that can be used to position any case in the 2-D space. The first dimension selects those

[1] A potential disadvantage of this is that zero means that a feature is not present, so partial descriptions are problematic.

high weighted features that account for the highest proportion of the variability and the second dimension selects less strongly weighted features.

A property of the analysis is that disease contours can be superimposed on the 2-D plot indicating degrees of typicality for cases of each disease. A case near the centre of the contours is highly typical of the disease whereas cases nearer to the periphery are less typical. A further property of the plots is that the proximity of two cases of a particular disease in the plot, i.e. their perceptual proximity, indicates the similarity of the two descriptions in the original multi-dimensional space, see Fig. 1 and also Section 4.2. The overview space has the property that the same perceptual distance between cases represents an increasing degree of similarity as one moves out from the centre of typicality, i.e. this matches the psychological finding that people can make finer similarity discriminations for more typically encountered cases.

In displaying cases for many diseases we adopt a largely hierarchical approach exploiting the "small worlds" metaphor [3]. We divide the diseases up into "small worlds" corresponding to small sets of confusable diseases, and compute separate composite weightings for each small world.

At present, subdivision of diseases into small worlds is based on the opinion of a single expert, but empirical work is in progress to verify these choices. Having computed the MCA weightings for the diseases in a particular small world, we can then use the MCA analysis to compute the separate likelihood contours for each disease in the chosen small world, see Fig. 1. The small world shown involves two broad categories of lesion. For more expert users the small world would need to be at a finer level of diagnostic discrimination.

By repeating this analysis for several small worlds, we have a set of possibilities against which a new case can be viewed. Just as a single small world can be displayed as a single overview plot, so a set of small worlds can be displayed in a composite form which presents the spatial relationship of one small world with another. Some distortion of the overall space may be needed to allow zooming in and out to visualize from the best viewpoint both the relationship between small worlds as well as the relationship between diseases within a small world.

3.3 Decision Support Methodology

Experts rapidly home in on a small world of possible diagnoses that explain most of the data; their visual and diagnostic reasoning are deeply intertwined and they try to reconcile clinical and case history information with data in the images after an initial detailed viewing of the images. In accordance with this view of radiological expertise, the following decision support methodology can be applied (with minor variations) whether the system is acting in the mode of "tutor" and offering a trainee an analysed case from the archive to diagnose, or whether the user (possibly more expert) is attempting to diagnose a case that is unknown to the system, essentially by comparing it to the others in the archive.

View MR case images. The first stage is to view and window the set(s) of image slices. The set is shown in the top left of Fig. 1. If case notes are available, e.g. for a case in the archive, these will not be accessible in the training system at

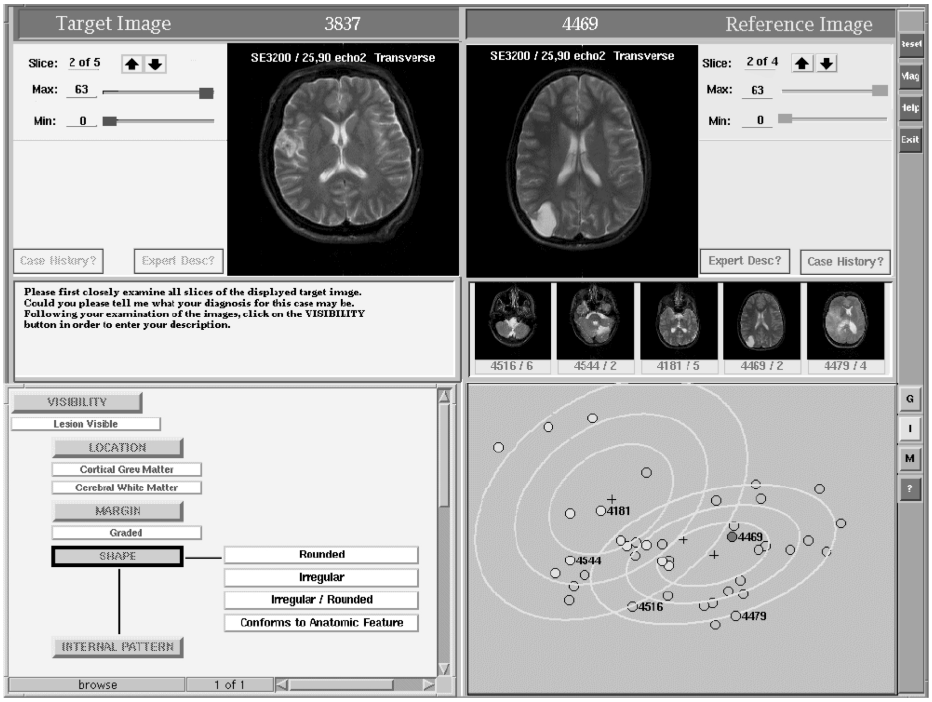

Fig. 1. The image description tutor. The small world of Glioma and Infarct is shown on the bottom right. Infarct cases are shown as lighter coloured dots. A partially completed image description is shown on the bottom left.

this point, so as to reinforce the primacy of the image over other data. Whether they should remain inaccessible, if available, in a decision support system for use by experts is a point of debate that careful user trials will help resolve.

Select and view small world from menu. In the current version of the system the overview plot can display a single small world (see Fig. 1), such as Glioma and Infarct, chosen from a set of small world possibilities. The rapid initial selection of hypotheses is accommodated by the user selecting and clicking on a single button to bring the chosen small world into the overview plot, see the bottom right of Fig. 1. We may have to enable the possibility of the user choosing more than a single small world at this stage, especially if it turns out that some diseases occur in more than a single small world.

Compare related cases from database. The overview plot is populated by dots, each dot representing a case from the archive. These dots are mouse sensitive and can be clicked on to bring the set(s) of image slices up on the screen, (see the top right of Fig. 1). The user can visually compare the images for cases from the archive with the case under examination. Moreover the position of the dot in the overview plot indicates, through its distance from the centroid for a particular disease, how typical that case is in comparison to the population for that disease. For example, Case No. 4161 in Fig. 1 indicates that, on the basis of how the images for this case appear, this is a very typical Infarct and similarly that case No. 4469 is a very typical Glioma.

Read clinical and presentation data. At this point it is important that the radiologist takes all the available data/information into account, if s/he has not already done so. Where a strict regime of delaying access to this data is in operation, access is now permitted to the clinical history and other case data if available.

If case is no longer problematic then exit. The images and the case data may render the case in hand unambiguous and at this point the user can exit, without further action other than to discriminate between the diseases in the selected small world if s/he can. In a future implementation, where the user is a trainee, and when the case under examination is known to the system, a reflective follow-up dialogue that is sensitive to the trainee's history of interaction with the system, the accuracy of their final diagnostic choice(s) and the process they went through as far as this is available to the system will be initiated e.g. which small worlds they explored in the overview plot, which cases within those worlds they called up in the comparison process and the manner in which they explored the images for the case in hand.

Describe case to system. Where the case is more problematic, either because of the trainee's lack of experience or because of its inherent difficulty, the user can engage in the additional task of describing the appearance of the lesion(s) on different sequences using the menu-driven structured image description language, see bottom left of Fig. 1.

See where the description lies in the small world. Using the same coefficients derived from the MCA analysis that produced the small world plot in the overview plot, the position of the case described by the trainee can be shown in the overview plot. Nearby and distant cases can then be examined by clicking on them to examine points of similarity and difference.

Check whether any other small world offers competing possibilities. At present there is only a single small world implemented, so the following steps represent future work. It may be that the position of the dot representing the current case lies in a region of difficulty such that it is either far outside the range of typicality for any of the diseases in the small world, or in a region equidistant from two or more disease centres.

In the former case, the user can investigate other small worlds to see if there are any which are both plausible, given what is known about the case, and display the dot for the current case nearer a disease centroid.

In the latter case, the ambiguous case, the system can offer advice as to which parts of the description have led to diagnostic uncertainty and/or to which further tests might be employed to reduce ambiguity.

Read off relative likelihoods from chosen small world. When the user agrees the description for the case under consideration and the best fitting small world is in view in the overview plot, then the relative likelihoods of the different diseases can be inferred from the position of the case relative to the disease contours of the different diseases in the small world. If required the system can compute diagnostic probabilities and display these to the user.

4 Evaluation

4.1 Description Language

The analytical power of the IDL has been partly tested by its application to the differentiation of multiple sclerosis (MS) from vascular disease [5] and the effects of HIV infection on the brain [6] Further insights into the predictive power of combinations of features will emerge as part of the continuing statistical analysis of the data, including the application of Multiple Correspondence Analysis.

4.2 Overview Plot

We have conducted a limited evaluation of the overview plot, based on the display of cases for a single disease. The evaluation [10] was carried out to investigate whether the statistically derived measures of typicality and similarity presented in the overview plot match the typicality and similarity judgements of radiologists.

A total of seventeen subjects took part in the experiment. These comprised four novices (with no knowledge of radiology), nine intermediates (4th year medics and radiographers with some knowledge of anatomy and imaging) and four expert neuroradiologists. The subjects were presented with the overview plot for a single disease, Glioma, on a computer screen with six cases removed. They were asked to fully explore all the presented cases by clicking on the points to bring up case images and associated descriptions. They were then shown the images and descriptions of each of the six cases previously removed and asked to place a marker representing each case at an appropriate position in the overview plot. Scores were derived for the similarity of each of the six test cases to all the other cases by computing their scaled Euclidean distances from the other points.

An ANOVA of the log distances showed significant differences between the novice, intermediate and expert placement of the cases in the overview plot ($F2, 60 = 3.150 for P < .05$). The average degree of agreement between human and MCA placement was in the expected order of expert (0.97), intermediate (0.95), novice (0.94).

Interviews with the subjects based on a structured questionnaire indicated that they found the overview plot easy to use and acceptable as a means for retrieving cases from the image archive. The evaluation suggests that the overview plot can provide a useful teaching device, to assist a trainee in forming a mental representation of the distribution of cases of a disease comparable to that of an expert.

5 Conclusion

We have described the main components of a largely visual decision support training system derived from an existing system to teach MR image description. This is based on implementing the notion of a small world as an interactive overview plot, based on an MCA analysis of cases from an archive. At present the components described are being re-implemented in Java to improve their portability and their modularity[2].

Much work is yet to be done. This includes choosing in a principled way the small worlds and including within the system some knowledge of the cues that evoke them; evaluating the decision-making leverage (if any) provided by the system, and evaluating the training potential of the system.

Acknowledgements

The work has been supported by ESRC Cognitive Engineering grant L127251035 and by EPSRC Medical Informatics grants GR/L53588 and GR/L94598. The original system was developed in POPLOG.

References

1. R. Azevedo, S. Lajoie, M. Desaulniers, D. Fleiszer, and P. Bret. RadTutor: The theoretical and empirical basis for the design of a mammography interpretation tutor. In *Artificial Intelligence in Education: Knowledge and Media in Learning Systems, Proceedings of AI-ED97*, pages 386–393. IOS Press, Amsterdam, 1997.
2. R. Azevedo and S. P. Lajoie. The cognitive basis for the design of a mammography interpretation tutor. *International Journal of Artificial Intelligence in Education*, 9:32–44, 1998.
3. C. Bishop and M. Tipping. A hierarchical latent variable model for data visualisation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):281–293, 1998.
4. G. du Boulay, B. Teather, D. Teather, M. Higgott, and N. Jeffery. Standard terminology for MR image description. In M. Takahashi, Y. Korogi, and I. Moseley, editors, *XV Symposium Neuroradiologicum*, pages 32–34. Springer, 1994.
5. G. du Boulay, B. Teather, D. Teather, N. Jeffery, M. Higgott, and D. Plummer. Discriminating multiple sclerosis from other diseases of similar presentation - can a formal description approach help? *Rivista di Neuroradiologia*, 20(7):37–45, 1994.
6. G. du Boulay, B. Teather, D. Teather, C. Santosh, and J. Best. Structured reporting of MRI of the head in HIV. *Neuroradiology*, 37:144, 1994.
7. M. Greenacre. *Correspondence Analysis in Practice.* Academic Press, London, 1993.
8. F. Hassebrock and M. Pretula. Autobiographical memory in medical problem solving. In *Proceedings of the American Educational Research Association Meeting*, Boston, Massachusetts, 1990.

[2] by Fernando De Andres Garcia at De Montfort University.

9. A. Hillard, M.-W. M., W. Johnson, and B. Baxter. The development of radiologic schemata through training and experience: A preliminary communication. *Investigative Radiology*, 18(4):422–425, 1985.
10. N. P. Jeffery. *Computer Assisted Tutoring in Radiology.* PhD thesis, De Montfort University, Leicester, 1997.
11. G. Joseph and V. Patel. Domain knowledge and hypothesis generation in diagnostic reasoning. *Medical Decision Making*, 10:31–46, 1990.
12. A. Kushniruk, V. Patel, and A. Marley. Small worlds and medical expertise: Implications for medical cognition and knowledge engineering. *International Journal of Medical Informatics*, 49:255–271, 1998.
13. A. Lesgold, H. Rubinson, R. Feltovich, P.and Glaser, D. Klopfer, and Y. Wang. Expertise in a complex skill: Diagnosing X-ray pictures. In M. Chi, R. Glaser, and M. Farr, editors, *The Nature of Expertise*, pages 311–342. Erlbaum, Hillsdale, NJ, 1988.
14. R. Macura, K. Macura, V. Toro, E. Binet, and J. Trueblood. Case-based tutor for radiology. In H. Lemke, K. Inamura, C. Jaffe, and R. Felix, editors, *Computer Assisted Radiology, Proceedings of the International Symposium CAR'93*, pages 583–588. Springer-Verlag, 1993.
15. G. Norman, L. Brooks, C. Coblentz, and C. Babcook. The correlation of feature identification and category judgments in diagnostic radiology. *Memory & Cognition*, 20(4):344–355, 1992.
16. V. Patel, G. Groen, and C. Fredicson. Differences between students and physicians in memory for clinical cases. *Medical Education*, 20:3–9, 1986.
17. E. Rogers. VIA-RAD: A blackboard-based system for diagnostic radiology. *Artificial Intelligence in Medicine*, 7(4):343–360, 1995.
18. H. G. Schmidt and H. P. Boshuizen. On acquiring expertise in medicine. *Educational Psychology Review*, 5(3):205–221, 1993.
19. M. Sharples, B. du Boulay, D. Teather, B. Teather, N. Jeffery, and G. du Boulay. The MR tutor: Computer-based training and professional practice. In *Proceedings of World Conference on Artificial Intelligence and Education (AI-ED 95), Washington, USA*, pages 429–436. 1995.
20. M. Sharples, N. Jeffery, D. Teather, B. Teather, and G. du Boulay. A socio-cognitive engineering approach to the development of a knowledge-based training system for neuroradiology. In *Artificial Intelligence in Education: Knowledge and Media in Learning Systems, Proceedings of AI-ED97*, pages 4/02–409. IOS Press, Amsterdam, 1997.
21. D. Teather, B. Morton, G. du Boulay, K. Wills, D. Plummer, and P. Innocent. Computer assistance for C.T. scan interpretation and cerebral disease diagnosis. *Neuroradiology*, 30:511–517, 1985.
22. D. Teather, B. Teather, K. Wills, G. du Boulay, D. Plummer, I. Isherwood, and A. Gholkar. Initial findings in the computer-aided diagnosis of cerebral tumours using C.T. scan results. *British Journal of Radiology*, 54:948–954, 1988.