Current Developments in the MR Tutor

Benedict du Boulay†   Rigoberto Bacallao Corrales‡   Lisa Cuthbert†,

Fernando De Andres Garcia‡,   George du Boulay⋆,   Nathan Jeffery‡,

Mike Sharples∗,   Briony Teather‡

Derek Teather‡

† *School of Cognitive and Computing Sciences*

University of Sussex, UK

Email: bend@cogs.susx.ac.uk

‡*Department of Medical Statistics, De Montfort University, Leicester*

⋆*Institute of Neurology, London*

∗*School of Electronic and Electrical Engineering, University of Birmingham*

Abstract: We have developed a prototype system, the MR Tutor, that aims to help trainees learn a systematic method of describing MR brain images by means of a structured image description language (IDL) [2]. The training system makes use of an archive of cases collected for research purposes and 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 some of methodological and practical issues which we are currently investigating, in particular, the refinement of the image description language as a trainee increases in expertise, and provision of support for diagnostic discrimination as well as image description.

1   Introduction

We have developed an image description training system [8, 9] 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 by an expert neuroradiologist (G. du Boulay) using the terms of the IDL.

The MR Tutor currently exists as a number of different prototypes. The main prototype, now implemented in Java, supports and scaffolds the process of describing single-sequence sets of images for single lesions. Another prototype provides the facility to adjust the level of support for the learner and to choose a case for tutoring based on the typicality and the trainee's performance so far. Depending on the expertise of the trainee it either produces immediate feedback to a description or provides an overall critique at the end [10].

One theme in the work is to provide the trainee with a new "way of seeing" the distribution of cases across diseases and to provide a largely visual means of making a differential diagnosis. Another theme of the work is to construct a training system that will be effective given the constraints under which neuroradiology is normally learned by apprenticeship, with the constraints of working within a busy hospital environment.

This work is part of a wider project to provide support and training in radiology that emphasises the value of systematic and careful analysis of images as the primary source of data *prior* to the attempt to reconcile observations with the clinical history and patient data. A further general aim of the work is to offer a *standard* language for describing MR images that should be of value to experts as well as trainees in that it will allow one expert's findings to be more readily compared with another's.

1.1 Image Description Language

The IDL describes the position and appearance of the features visible in 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.

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 with definitions and indicative examples 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.

1.2 Overview Plot

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.

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 value. Multiple Correspondence Analysis (MCA) is a statistical technique for data reduction and visualisation [4]. 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.

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 high weighted features that account for the highest proportion of the variability and the second dimension selects the less strongly weighted combination of features.
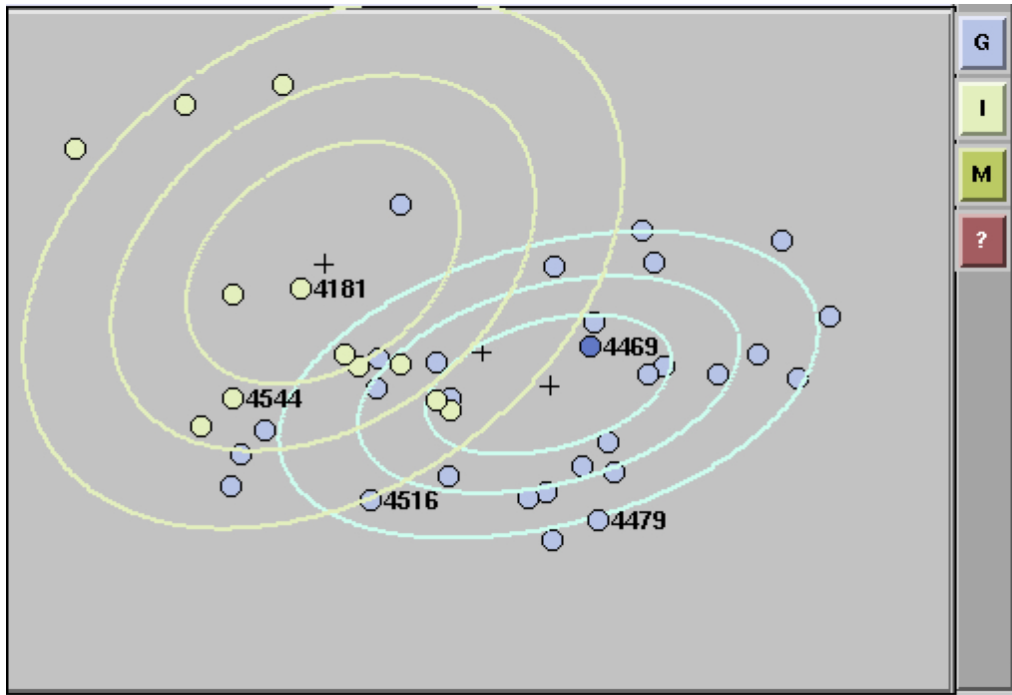
Figure 1: The small world of Glioma and Infarct

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 (it has feature values in common with many other cases of the disease) whereas cases nearer to the periphery are less typical. A further property of the plots is that the proximity of two points indicating 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. The typicality contours match the psychological finding that people can make finer similarity discriminations for more typically encountered cases.

1.3 Small Worlds

Once the trainee has learned how to use the image description language to describe and compare cases within the context of a single disease, it is necessary to provide support for description and diagnosis in the context of confusable diseases.

Medical experts possess highly structured knowledge that provides various kinds of short-cut to the small set of hypotheses that need to be considered in any situation. When presented with a new case they rapidly home in on a number of "critical cues" that guide them to consider a small set of possible hypotheses of possible diseases ("small world") [6, 7].

In the tutor we are dividing the diseases up into small worlds (e.g. two, three or four members each), corresponding to sets of confusable diseases, and compute separate composite weightings for each small world, see Fig. 1.

1.4 Organization of the Paper

At present the MR Tutor is being developed in a number of ways, in particular: (i) to give support for diagnosing as well as describing, based on the diagnostic notion of a small world [1], (ii) to improve the interface in line with the results of an earlier workplace study, (iii) to provide better tutorial support both for the *processes* of describing and diagnosing as well as

4

for the quality of *outcomes* of description and diagnosis, and (iv) to find ways of assisting the expert to describe images for the archive by providing means of semi-automatically analysing images, e.g. determining the nature of the lesion boundary, given that its location has been found by hand.

This paper concentrates on two of the methodological and practical issues which we are currently investigating. These are, first, the relationship between the simplified and the full image description language and, second, support for diagnosis as well as description. The main emphasis here is on how the system can adjust the complexity of the task and its support as trainees grow in expertise.

2 Adjusting Task Complexity

2.1 Simplified and Full Description Language

The prototype tutor has been developed using a simplified version of the description language requiring some 30 lesion features to be described. Two further simplifications have been made thus far: (i) to restrict the cases to those involving a single lesion, and (ii) to use images from a single sequence or echo for each case. The full, "expert" version of the description language includes 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 [2].

Completing such a detailed description is hard work and time-consuming, especially where multiple sequences/echoes and/or multiple lesions are involved — a description of the largest of each lesion "type" is required for each sequence/echo. Expending such effort makes sense in the context of a research project but is harder to justify in day-to-day work, and to some extent in apprenticeship-style training based on day-to-day work. So it is important to increase the complexity carefully when moving from the simplified towards the full image description language.

Uncovering more of the complexity of the language as the trainee becomes more expert will have an effect on how the overview plots look. Changing the list of features used in descriptions will require a recomputation of all weights in the MCA. This may have the effect of changing the distribution of cases in the overview plot as different features may now account for the variability among cases. So the trainee may not only be faced with an evolving language but also an evolving set of overview plots.

2.2 Multiple Sequences/Echoes

The coordinates of a case in the overview plot, either for a single disease or for a small world, are currently computed based on a single sequence (i.e. T2 weighted) and a case is represented by a single point.

The distribution of points of a particular disease, glioma (say), in the overview plot depends on whether the overview plot is constructed just for that disease (i.e. glioma) or for that disease plus others (e.g. a small world). The reason that the shape of the distribution may change is that the MCA will compute a different set of weights to construct the best fitting 2D subspace for the cases of glioma on their own compared to glioma plus some other disease(s). As a key aspect of training is to understand how different image features contribute to discrimination between diseases in a small world, it follows that the overview plot must be

constructed from the cases of the appropriate small world since this MCA plot will be based on the linear combination of feature values that best spreads the cases in the small world.

Most of the cases in our archive have images from several sequences/echoes (typically two or three) and multiple sequences/echos raise a further issue for the overview plot representation. When more than one sequence or echo is available we could provide an overview plot for each sequence/echo or a single overview plot across all sequences/echoes. There may be some value in presenting a view of the diseases on each sequence/echo to illustrate how different sequences/echoes provide more or less separation of the diseases. However this presents the trainee with a difficult additional task — that of integrating information from several overview plots. We are currently extending the methodology so that overview plots for each sequence/echo may be viewed together with an overview plot constructed from a principal components analysis (PCA) of the MCA results from each sequence/echo. PCA can be used at this stage because the MCA results are weights expressed in real numbers.

3 Support for Diagnosis

3.1 Choosing Small Worlds

We are currently conducting experiments with expert radiologists to determine plausible small worlds to include in the system. This experiment involves asking radiologists to suggest differential diagnoses that are are hard to discriminate (and also to indicate the frequency of occurrence), as well as to comment on suggestions for small worlds made by our own expert (G. du Boulay). In terms of developing the tutor we will be constrained initially by the distribution of cases available in the archive — we may not have enough examples of some diseases to populate all the small worlds suggested — but the study will indicate for which diseases further cases and their descriptions will be needed. The frequency of occurrence information will be helpful in determining the relative population sizes of the diseases in the tutor to reflect case mixes at different kinds of hospital.

In the longer term it may prove useful to allow the trainee *additionally* to "assemble" a small world from individual diseases. At present the MCA is conducted off-line and its results hand-coded into the prototype tutor. Allowing an arbitrary small world to be assembled would require the MCA to be computed for that particular small world, as storing the MCA weights for every possible combination of diseases would be too costly in storage terms. However running an MCA analysis is computationally costly and requires some statistical validation, and its not yet clear whether this would impose unacceptable time delays in the tutor's response.

3.2 Diagnostic Discrimination

In [1] we describe a method of helping the trainee to use their visual intuition to first home in on an initially plausible small world, and then to refine that choice in the light of systematic analysis of all the available images and any other available data. It will be straightforward to check on the degree to which the trainee follows this diagnostic methodology, has examined all the slices (in all the sequences) and made reasonable use of the overview plot to call up and examine potentially similar cases.

An important part of the process of discrimination is reconciling the possibilities suggested by the (primary) image data with the (secondary) information in the case notes. At present the knowledge representation of the system is based predominantly on the image description language and thus on what the lesions look like, where they are located and how they affect nearby areas of the brain. Although clinical and other patient data is available (and can be

viewed by the trainee at certain points in the process), this data has not yet been incorporated into the knowledge representation of the system. Doing this could enable the development of the system to comment on the trainee's beliefs about confirming and disconfirming evidence in the secondary data.

While many of our small worlds will contain diseases that are genuinely hard to discriminate, further analysis of our archive is likely to provide evidence of those "critical cues" that may be helpful in particular discriminations [3]. These cues can then be incorporated explicitly into the tutoring strategy.

Another area of development would consist of advice about what further image data might usefully be collected to refine a particular discrimination. This was provided in an earlier system, BRAINS, for CT scans [11, 12].

3.3 From the Overview Plot to Descriptions

Once the trainee has made a description of a case, that description can be displayed as point in the overview plot. When such a point lies in a region of overlap between diseases, it may be useful to determine which parts of the description would need to change in order for the point to move nearer or further away from a particular disease's centre of typicality.

The mapping from descriptions to coordinates is many-one. Each description yields a unique point in the overview plot under MCA analysis. But because it is a projection in multi-dimensional space onto two dimensions, a single point in the overview plot can be produced by many descriptions. Given a space that is densely populated by cases, and making an assumption that the space is smooth one can interpolate a small number of plausible case descriptions for a particular point. This enables the computation of which parts of the description would need to change in order for the point to move nearer or further away from a particular disease's centre of typicality.

3.4 Partial Descriptions

In some situations it may be helpful (and save time) for the trainee to enter only a partial description of a case, e.g. concentrating on location and shape of the lesion only. However, one of the acknowledged problems with MCA is the handling of missing data. It is not appropriate simply to equate absence of data on a feature value with a feature value known to be absent (code '0'). This raises a difficulty in the representation of partial descriptions where only some feature values are specified, in that it is not obvious how to code the remaining features. Partial descriptions would be very useful in that they would (in principle) allow a quick entry of data to get a sense of where a case lies in the disease spaces, followed by more detailed input of certain parts of the description if that looked a sensible course of action.

One possibility might be to represent a case with missing data as a *region* rather than as a point in the overview plot. Where the missing data consisted of feature values with small weights (i.e. features values that do not account for much of the variability among cases in that small world), such a region would be small. The converse is also true, so a partial description that omitted a highly weighted feature would cover a large region of the plot — thus indicating the importance of the missing data.

Another possibility might be to estimate the missing data heuristically, on a similar basis to the backward mapping described above.

3.5 Geometric and Topological Intuitions

We have already looked at the degree of accuracy with which novices, intermediates and experts are able to place new cases (given their descriptions) in a partially populated overview plot for a single disease. This study showed that experts placed new cases in a similar way to the MCA analysis — and were closer to the MCA than either intermediates of novices [5].

We are about to embark on a number of other studies of geometric and topological intuitions supported by the overview plot, in particular using plots of small worlds, rather than of single diseases. One such study will measure the extra diagnostic leverage (if any) provided by overview plot, especially how judgements of relative odds between diseases are influenced by the visual feedback of seeing the description of a new case as a point in the overview plot.

4 Conclusions

This paper has described current developments in the MR Tutor — a system that aims to help trainees learn a systematic method of describing and then diagnosing from MR brain images by means of a structured image description language. It has concentrated on adjusting the complexity of the task to match the increasing expertise of the trainee, especially via the relationship between the simplified and the full image description language. It has also discussed provision of support for diagnosis as well as description via enhancements to the overview plot.

The MR Tutor has recently been reimplemented in Java and is currently being extended with a more versatile interface for displaying image sequences. We are also exploiting more of the images from the archive within the tutor, exploring how best to incorporate images (and their descriptions) containing contrast enhancement, and investigating ways of constructing image descriptions semi-automatically.

5 Acknowledgements

References

[1] B. du Boulay, B. Teather, G. du Boulay, N. Jeffery, D. Teather, M. Sharples, and L. Cuthbert. From description to decision: Towards a decision support system for MR radiology of the brain. In W. Horn, Y. Shara, G. Lindberg, S. Andreassen, and J. Wyatt, editors, *Artificial Intelligence in Medicine: Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making, AIMDM99, Aalborg, Denmark. Lecture Notes in Artificial Intelligence 1620*, pages 93–102. Springer-Verlag, Berlin, 1999.

[2] 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.

[3] 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.

[4] M. Greenacre. *Correspondence Analysis in Practice*. Academic Press, London, 1993.

[5] N. P. Jeffery. *Computer Assisted Tutoring in Radiology*. PhD thesis, De Montfort University, Leicester, 1997.

[6] G. Joseph and V. Patel. Domain knowledge and hypothesis generation in diagnostic reasoning. *Medical Decision Making*, 10:31–46, 1990.

[7] 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.

[8] 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.

[9] 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.

[10] T. Simpson. The new MR tutor program: Improving and extending computer-based radiology tuition. Master's thesis, School of Cognitive and Computing Sciences, University of Sussex, 1998.

[11] 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.

[12] 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.