Clinical reasoning skill acquisition: identifying learning issues and developing vicarious learning resources.

Richard Cox†, Kirsty Hoben‡, Barbara Howarth#, John Lee§,
Julie Morris#, Jianxiong Pang†, Susen Rabold§ & Rosemary Varley‡

† Department of Informatics, University of Sussex
‡ Human Communication Sciences, University of Sheffield
# Education, Communication& Language Sciences, University of Newcastle
§ School of Informatics, University of Edinburgh

Vicarious learning and case-based teaching of clinical reasoning skills project

Abstract

We report results from studies of Speech and Language Therapy (SLT) students' diagnostic reasoning. The 'patients' were presented via PATSy, an extablished case-based interactive e-learning environment (www.patsy.ac.uk). Participant dyads were encouraged to discuss aloud their reasoning as they interacted with the virtual cases, made assessment decisions, generated hypotheses, etc.

Rich real-time process data was recorded - dialogues were videotaped, the computer display was dynamically captured, together with PATSy-generated student/system interaction logs which included reflective comments entered by the students during the sessions.

Qualitative analysis tools (NITE) were used to synchronise playback of the multiple data streams and to facilitate coding and markup of participant dialogues and learning difficulty events. Analyses were informed by a conceptual framework of clinical reasoning developed by the project team containing domain-specific and general reasoning components, together with a stage-model of diagnostic reasoning.

T·L·R·P TEACHING & LEARNING RESEARCH PROGRAMME
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Results indicate, *inter alia*, that learning difficulties arise from poor general reasoning skills (eg. inability to design an effective testing strategy) and/or poor domain-specific knowledge (eg. knowledge of assessment tests). These difficulties were evidenced by participants' discussions and confirmed by triangulation with other data sources.

We are currently moving into the second phase of the project. Vicarious learning (VL) resources are being developed using task-directed discussion (TDD) methodology. A range of TDD exercises are being designed to address the issues that were elucidated by the diagnostic reasoning studies. The VL resources will take the form of video clips of educational dialogue about the learning issues. Future students will be able to access and learn from them. The effectiveness of such 'learning by observing the learning of others' (aka VL) will be evaluated in the project's final phase.

Keywords: Vicarious learning, observational learning, collaborative learning, case-based teaching, clinical reasoning, diagnostic reasoning, e-learning, allied health professions, speech and language therapy, clinical training, clinical education, dialogue analysis, discourse analysis, protocol analysis, verbal analysis, qualitative methods, higher education, CPD , task-directed discussion, think-aloud protocol

Overview of project

The overall aim of the project is to produce and evaluate learning resources that will assist students acquiring clinical reasoning skills through vicarious learning. Sources of difficulty may arise from a lack of domain-specific knowledge and/or a lack of general reasoning skills. We aim to develop high quality learning resources which will enhance students' reasoning and evaluation skills. The learning resources will take the form of vicarious learning (VL) resources. VL has been described as 'learning by observing the learning of others' (Mayes, Dineen, McKendree & Lee, 2001; McKendree, Stenning, Mayes, Lee & Cox, 1998; Cox, McKendree, Tobin,Lee & Mayes, 1999).

The project has three phases and is currently in transition between phases one and two.

An established case-based e-learning system called the Patient Assessment Training System (PATSy) plays a central role in all phases of the project. Detailed information about PATSy can be found on the web at `www.patsy.ac.uk`, and in Lum & Cox (1998), Cox & Lum (2004) and Lum, Cox, Kilgour, Morris & Tobin (1999).

In the first phase of the project, the aim was to identify learning difficulties in clinical reasoning, where this was defined as 'an inability to apply formal knowledge in a clinical setting'. The PATSy system was used as a tool to investigate Speech and Language Therapy students clinical reasoning in this phase. Findings from the phase 1 studies are summarised in the first section of the paper.

In the second phase of the project, we are developing vicarious learning resources that address the learning issues identified in phase 1. We are using 'task directed discussion'

(TDD) methodology as a basis for generating useful and re-usable educational dialogues. The dialogue participants will be pairs of students. This phase of our work is described in the second section of the paper.

Over the course of phase two of the project, the video clips of the dialogues (ie. the VL resources) will be incorporated into PATSy. The system will be extended to include a database of VL resources with each item in the database 'tagged' with metadata describing the case(s) it relates to, the topic(s) that it addresses, the dialogue participants, etc. The PATSy system will be extended so that it is capable of offering appropriate VL resources on (student) demand in situations where the user runs into a reasoning impasse.

The third phase of the project will consist of a controlled study designed to assess the efficacy of VL as a mode of learning. We also aim to distill our findings into the form of 'how to' guides for educators in other subject areas to use so that they can develop effective VL resources for use in their domains.

Phase 1 Clinical reasoning studies

The aim was to identify learning difficulties in clinical reasoning, where this was defined as 'an inability to apply formal knowledge in a clinical setting'.

Seventeen pairs of students (dyads) were studied as they diagnosed previously unseen language disorder cases on PATSy. The participants were undergraduate and postgraduate students of SLT at Newcastle and Sheffield universities. Each pair diagnosed one patient.

The participants were randomly assigned to 1 of 3 cases (2 adult, 1 paediatric). Fairly complex cases were deliberately chosen in order to elicit difficulties in reasoning. Each dyad spent one hour working towards a diagnosis of the patient's communication difficulties.

Participants were instructed to:

> ...assess the communication skills of a [virtual] patient in order to produce a set of statements that describe key impairments shown by the case, and if you can, an overall diagnostic category, such as phonological delay, agrammatic aphasia...

Several sources of rich process data were collected: student-system interactions (logged by the PATSy system automatically); video of the student dyads engaged in dialogue during collaborative diagnostic reasoning, notes entered into the PATSy log collaboratively by the participants (Figure 1), and real-time video capture of the computer screen[1].

The video data sources were integrated and synchronised at playback using the NITE tools[2] developed at Edinburgh. A screenshot of NITE in use is shown in Figure 2.

[1] Using 'TechSmith Camtasia Studio' dynamic screen capture software, see - `www.techsmith.com/products/`
[2] `www.ltg.ed.ac.uk/NITE/`

Participant RS (top of page)

- **Access**: Patient RS video clip 1 in QuickTime 3 format [Thu Jun 9 11:46:19, 2005]

User Notes (at Thu Jun 9 11:50:08 BST 2005):

  - **Conclusions?**: wordfinding difficulties, jargon-y dysarthric speech

- **Access**: Medical History (patient RS) [Thu Jun 9 11:50:14, 2005]
- **Test**: ADA S2 Wr: Word-Picture Matching (written) [Thu Jun 9 11:51:49, 2005]
- **Test**: ADA S2: Word-Picture Matching (spoken) [Thu Jun 9 11:52:06, 2005]
- **Test**: Cinderella Sample (spoken) [Thu Jun 9 11:52:31, 2005]

User Notes (at Thu Jun 9 11:58:02 BST 2005):

  - **Conclusions?**: incoherent, word finding, jargon, content words ommitted, empty speech
  - **Where next?**:more test data, specifically

- **Test**: TROG (spoken) [Thu Jun 9 11:58:23, 2005]
- **Test**: Maximal Pairs [Thu Jun 9 11:58:38, 2005]

User Notes (at Thu Jun 9 12:01:44 BST 2005):

  - **Conclusions?**: Trog score very poor, audtitory discrimination poor at maximal pair task - impact on TRog score
  - **Where next?**:hearing, picture verification

- **Test**: Morris 50 Item Picture Verification Test: Session 1 [Thu Jun 9 12:02:22, 2005]
- **Test**: CAT 08 - The Cognitive Screen - Arithmetic [Thu Jun 9 12:06:30, 2005]

User Notes (at Thu Jun 9 12:08:00 BST 2005):

  - **Conclusions?**: repetition poor unsurprising given auditory discrimination naming poor, worse for verbs

User Notes (at Thu Jun 9 12:08:32 BST 2005):

  - **Conclusions?**: written and spoken word comp about equal

- **Test**: CAT 27 - The Language Battery - Written Picture Naming [Thu Jun 9 12:09:11, 2005]

User Notes (at Thu Jun 9 12:09:56 BST 2005):

  - **Conclusions?**: written picture naming worse than spoken picture naming

- **Test**: CAT 19 - The Language Battery - Object Picture Naming [Thu Jun 9 12:10:34, 2005]

User Notes (at Thu Jun 9 12:12:32 BST 2005):

Figure 1: Example of PATSy log data - students' notes (lines beginning with 'User Notes') are interleaved with machine-generated logs of PATSy items accessed during the session (lines beginning with 'Access' & 'Test'). This student dyad are working through an adult case (RS).

The data from selected dyads have been analysed using a general conceptual framework that characterises clinical reasoning as a complex skill comprising domain specific knowledge and general reasoning skills (Howarth, Hoben, Varley, Lee & Cox, 2005)[3].

The framework (Figure 3) is derived from research on scientific reasoning (eg. Klahr, Fay & Dunbar, 1993), models of general clinical reasoning (Hayes & Adams, 2000) and empirical studies of clinical reasoning (Patel, Groen & Fredericson, 1986; Cox & Lum, 2004).

[3]Paper available at `www.tlrp.org/dspace/retrieve/318/Cox+full+paper+conf+04+revised.doc`

Figure 2: NITE tools used to synchronise playback of videos of student dyad and computer screen activity recordings (nb. participants' faces have been obscured for reasons of ethics and confidentiality).

Separate coding systems have been developed for coding a) impasses in problem solving (general reasoning) activity and b) students' diagnostic statements (eg. relating to hypotheses).

Within the context of insight theory (Ohlsson, 1992), an impasse in reasoning is viewed as a state of mind in which the problem solver has exhausted all possibilities and cannot think of a way forward. Behaviourally, impasses are characterised by the cessation of problem-solving activity. Fleck and Weisberg (2004) applied their scheme to verbal protocols of participants performing a problem solving task with a well-defined initial state and an ill-defined goal state, namely Dunker's candle problem (Dunker, 1945). Although assessing patients' communication skills is a much more complex task, subjective observation of participants' discussions suggested certain parallels. For example, some participants indicated that they were out of ideas or were confused about the problem, or more specifically where a line of reasoning was going. Thus the scheme was adapted to reflect the more complex nature of the current problem. It consists of 7 types of behaviour. Examples include:

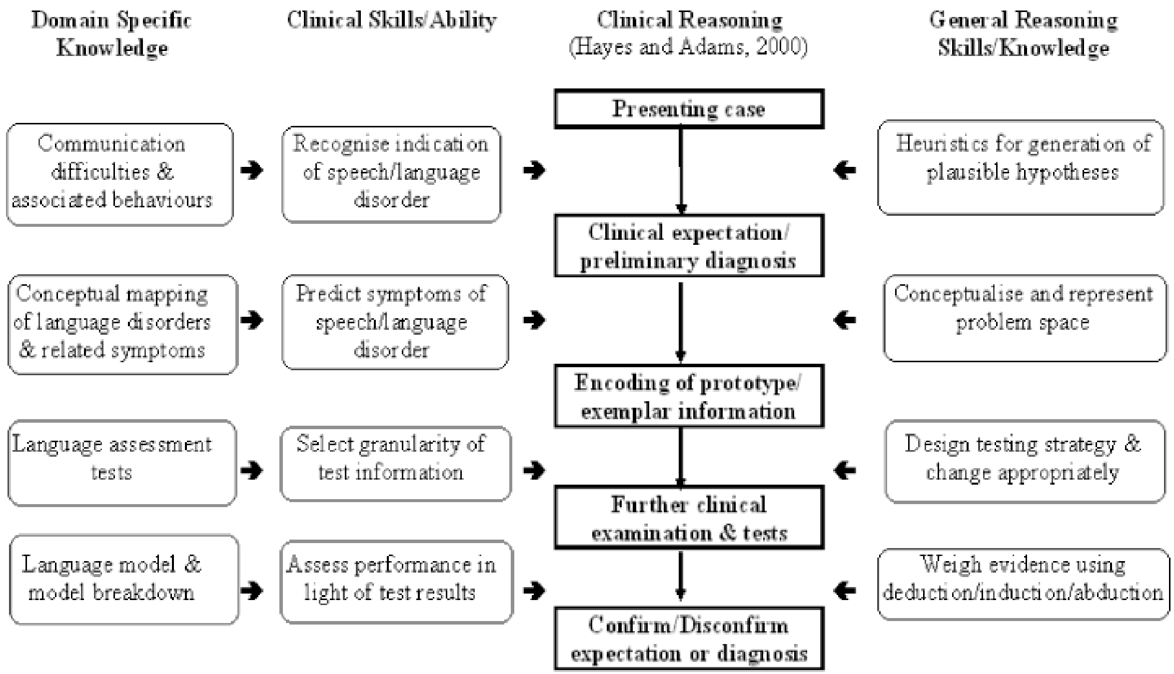

Figure 3: Some of the components of the complex skill of clinical reasoning. The diagram depicts stages involved in process of assessing patient and assigning a diagnostic category and includes general reasoning and domain-specific knowledge components.

- Rereading the problem (i.e. task goal on instruction sheet) three or more times in succession (type 2)
- Guessing at the solution (eg. stating a diagnosis 'out of the blue'). Expressing a lack of confidence in ability to generate a solution (type 4)
- Demonstrating fixation on the inability to understand why a prior solution (hypothesis) should be rejected (type 6)

The second scheme is a six-stage model of diagnostic statements which is being used to characterise the sophistication of students' clinical observations and to ascertain whether their hypotheses are actually testable using the information on the PATSy system. Examples from this scheme include:

- Noting information without making any judgement (eg. '24 out of 40 on that test') (level 1)
- Concrete clinical observation - making a judgement from an observation but not extrapolating from it or using professional register in reference to it (eg. 'he found that difficult') (level 2)

- True hypothesis which is couched as a predictive statement (eg. 'Could it be something like Asperger's syndrome?') (level 4)
- Diagnostic certainty (eg. 'autism', 'deficits in the central semantic system') (level 5)

Both schemes are currently undergoing iterative development and are being refined to ensure acceptable levels of inter-rater reliability.

Clinical reasoning study results

As mentioned earlier, sources of students' difficulty may arise from a lack of domain-specific knowledge and/or a lack of general reasoning skills. One kind of error derives from a lack of speech and language therapy domain-specific knowledge (eg. of communication disorders, knowledge of diagnostic tests and what they measure). Another type of impasse or error derives from poor general reasoning skills (eg. ability to use efficient search heuristics early in testing).

Some of the general reasoning skills that we (eg. Howarth et al., 2005) identified include:

- knowledge of heuristics for the generation of hypotheses based on theory or observation
- ability to conceptualise the problem space and represent it appropriately in terms of size and form (eg network, hierarchy, matrix)
- ability to design a testing strategy (eg. vary one thing at a time (Klahr, 2000))
- change strategy as a function of reasoning stage (eg. do broad brush tests first)
- ability to weigh evidence

The results also suggest that students at an early stage of clinical reasoning skill acquisition tended to: 1. note information without evaluating its clinical significance, 2. be limited in the extrapolations they make from clinical observations, 3. generate true (well-formed) hypotheses rather infrequently and, 4. show only limited use of professional register in their discussions.

In contrast, students who perform better at clinical reasoning, and experts, tended to generate true hypotheses and demonstrate higher levels of certainty for a particular diagnosis and for ruling out alternative diagnoses. They also couched their judgments and clinical observations using terms and phrases from the SLT professional register (eg. the use of phrases such as "its a pragmatic disorder", "definitely autism", "deficits in the central semantic system").

Another finding was that some students' testing sequences were 'driven' to some extent by the order in which the PATSy system lists them on screen. This tendency was detected via the analysis of dialogue samples such as:

> Student 2: 'Right, OK, where shall we go next then? I reckon, shall we just go through them and see what, ooh. Right, okay.'
> Student 1: 'I think we just should work through in order really. All right, the things that are odd. Sentence repetition might be interesting. Shall we start from the top then?'

The PATSy system logs and computer screen recordings confirmed this pattern. Better performing students (and experts) tended to select broad, search-space reducing tests early in their reasoning and then move on to fine grained, problem-isolating, 'higher resolution' testing as they close in on their diagnosis.

The diagnostic statements coding scheme has also highlighted interesting differences between students at different stages of clinical reasoning skill acquisition, and between students and experts[4]. Figure 4 shows the proportion of types of diagnostic statement made in the PATSy log during a session. It shows that students at earlier stages of clinical reasoning skill acquisition have a lower proportion of high level diagnostic statements (that is, level three and above) regardless of how many statements they make[5].

A further finding was that students who were more diagnostically accurate tended to make higher level statements more of the time and progress towards level five diagnoses over the time course of the diagnostic session. In contrast, the average diagnostic level of less skilled dyads tended to oscillate between levels 1, 2 and 3 (with an average of level 2), and not improve over the course of the session.

Phase 1 data analyses and an extensive literature review led to the development of a model of clinical reasoning development (Figure 5). Though the model was developed, in this case, for speech and language therapy, it could also be applied to clinical reasoning in other health science contexts.

Development of VL resources using TDD methodology

We are using task-directed discussion (TDD) exercises as a methodology for generating pedagogically effective dialogues for use as vicarious learning resources. TDDs are language games of a type commonly used in ESL teaching and second language learning (Skehan, 1988). A TDD is a discrete language activity with a set goal (Mayes et al., 2001). The topics of the TDD exercises we are developing are informed by the results of the clinical reasoning studies conducted in Phase 1 of our project.

From the researcher's point of view, a TDD is a device for generating discussion where spontaneous discussion is infrequent and where capturing naturally occurring dialogue is

[4] Data from three SLT experts (experienced educator-clinicians) was available from a previous study (Cox & Lum, 2004)

[5] The expert completed the case over a longer time period and therefore the absolute number of statements made cannot be compared across student and expert.

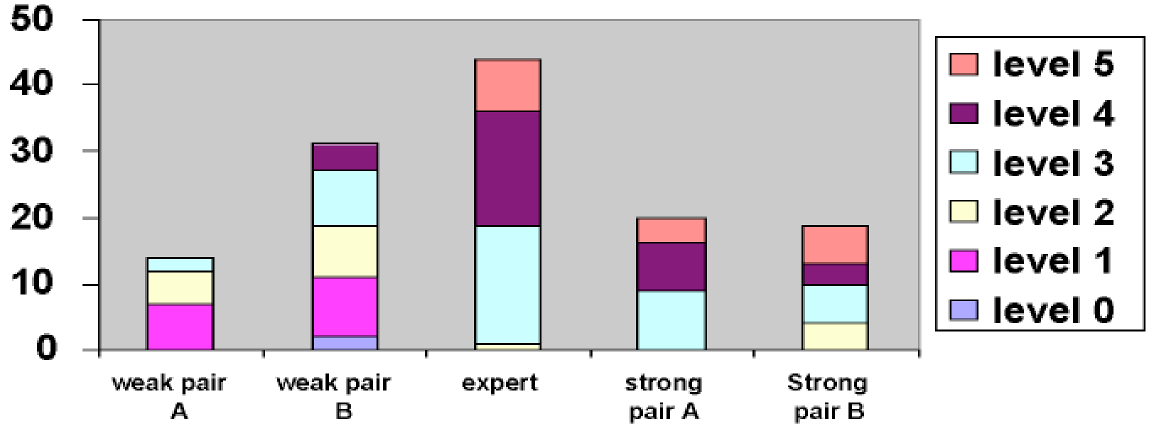

Figure 4: Composite column graph showing proportion of types of diagnostic statement made in the PATSy log during a session. Level 5 is diagnostic certainty, and level 1 is judgement-free noting of information only without an hypothesis. Y-axis is number of statements.

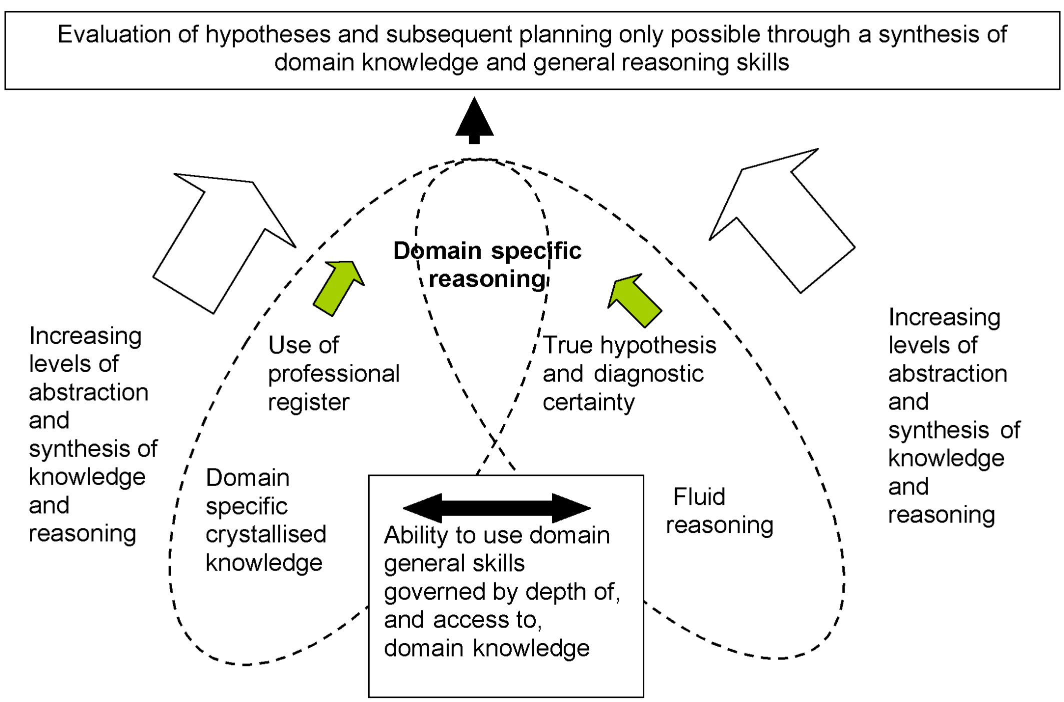

Figure 5: A model of clinical reasoning development in speech and language therapy.

difficult and uncontrollable.

There are numerous methodological advantages of the TDD methodology over 'natural' (eg. classroom-based) discussion. Diagnosing SLT cases is a high cognitive load task. Were we to try and capture spontaneous dialogue from pairs of students engaged in the full-blown clinical task it is likely that the potential for re-usability as VL resources would be quite low. Hence a first reason for using TDDs is to avoid over-taxing participants' cognitive resources to such an extent that there is no remaining capacity for dialogue or verbalisation.

In classroom and small-group tutorial contexts students are often reluctant to articulate the current state of their understanding. Any utterances they do make may be 'socially edited' to a large degree to mask the true extent of their (mis)understanding. The structured nature and limited (topic) scope of TDDs provide a means of limiting the potential for 'loss of face' in front of others. TDDs provide a context in which students are more likely to explain their thinking and identify areas in which they are uncertain. TDDs also encourage 'audience designed' dialogue via other-directed explaining (see Lee, Dineen, McKendree, & Mayes,(1999) and Cox (2005) for further discussion).

A typical domain-specific topic for a TDD might be designed to elicit discussion about the definition of a concept. (eg. *'Please provide a definition of paraphasia'*) or about how concepts compare (*'Discuss when the Graded Naming Test might be a better choice than the Nickels Naming Test when examining anomic difficulties.'*).

Other types of domain-specific TDD will cover topics such as knowledge of tests of language comprehension/production (eg. their names and what they measure), knowledge of models of normal language production and knowledge of language pathologies.

Some of the domain-specific TDD exercises will utilise diagrams as mediating representations. As Eraut (2000) has noted, a person's 'capability to tell' can be assisted by the presence of a mediating object such as a diagram. Suthers et al. (2001) report, on the basis of extensive research on collaborative tools for diagramming argument structure, that 'manipulable representations' primarily serve as 'a stimulus and guide for ... discourse...' (p. 13). We have therefore developed a computer-based interactive tree-diagram of the cognitive subprocesses associated with language comprehension and production (Figure 6)[6]

Further TDD examples focus on the use of professional register, eg. *"Read the following description of a client's communication difficulties and then rephrase it in professional terminology"*. Conversely, scientifically-phrased statements can be translated into everyday English. These kinds of activity should help to bridge the colloquial and scientific language registers and have been advocated in science education (Lemke, 1990).

Yet other types of TDD will address general reasoning skills *"Which of the following would be the best strategy when diagnosing a client's communication difficulty, and why?: a) Observe, Hypothesis, Test, Evaluate; b) Hypothesis, Observe, Test, Evaluate; or c) Test, Hypothesis, Test, Hypothesis"*).

[6] The interactive diagram will also be used as a knowledge elicitation tool with which SLT domain experts can provide meta-data for incorporation into the PATSy system.

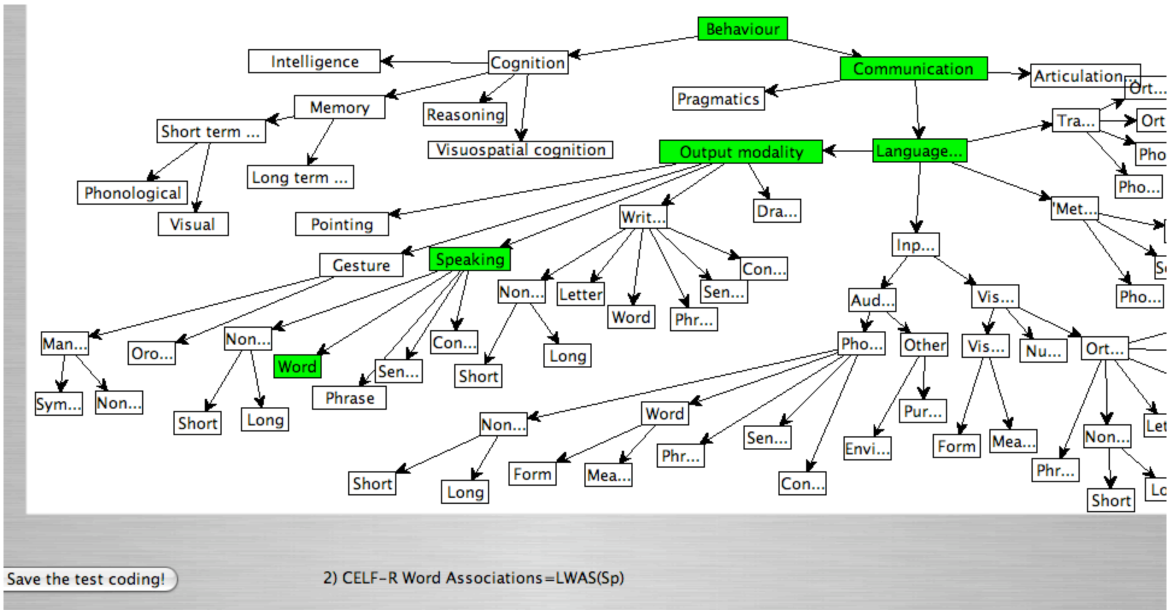

Figure 6: Interactive tree diagram to be used as a mediating representation between dialogue participants in some TDD exercises.

Conclusion and future work

As mentioned in the introduction, the TDD-derived dialogues will be recorded, edited and incorporated into a structured database for timely presentation by PATSy on demand from students when they strike a reasoning impasse. In phase III of the project we will use PATSy to evaluate the cognitive and social effects of the VL resources in a controlled experiment.

We also plan to increase PATSy's ability to adapt and respond to students. In order to 'know' what VL resources to offer a student, the system will need to be able to shadow the user and trace his or her clinical reasoning activity to some extent. It will be straightforward for the system to detect when students are merely working their way down the test list for a case. However, we would also like to build into PATSy a module in which student reasoning patterns can be stored and for this store to grow over time as more and more users interact with the system (i.e. a student model of the kind typically found in intelligent tutoring systems). We already have a considerable amount of data with which to 'seed' such a module.

A knowledge base of information about each test on the system, and what aspects of language, cognition, etc that it measures, may also be added. The knowledge base of information about tests, together with the student model, could provide a basis for tracing, in real time, the coherence of a learner's test selection patterns. Thus, if a student using PATSy encounters a reasoning impasse and requests a vicarious learning resource, the system will be capable of offering items that are a) relevant to the case being worked on, b) commensurate with her current stage of reasoning, and c) appropriate for the type of impasse that she is confronted with.

Acknowledgements The support of the Economic and Social Research Council is gratefully acknowledged (ESRC TLRP Phase III project grant RES-139-25-012).

References

Cox, R. (2005) Vicarious learning and case-based teaching of clinical reasoning skills. Pedagogy paper presented at the 'Technology, learning relationships & networks' seminar. Third seminar in the joint Centre for Research in Lifelong Learning (CRLL) and Economic & Social Research Council (ESRC) Teaching & Learning Research Programme (TLRP) Thematic Seminar Series on *Contexts, communities, networks: Mobilising learners' resources and relationships in different domains.*. Open University, Milton Keynes, October.

Cox, R., & Lum, C. (2004) Case-based teaching and clinical reasoning: Seeing how students think with PATSy. In Brumfitt, S. (Ed.) *Innovations in professional education for Speech and Language Therapists*, London: Whurr.

Cox, R., McKendree, J., Tobin, R., Lee, J. & Mayes, T. (1999) Vicarious learning from dialogue and discourse: A controlled comparison. *Instructional Science*, **27(6)**, 431–457.

Duncker, K. (1945). On problem solving. *Psychology Monographs*, **58**, (Whole No. 270).

Eraut, M. (2000) Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, **70**, 113–136.

Fleck, J.I. & Weisberg, R.W. (2004) The use of verbal protocols as data: An analysis of insight in the candle problem. *Memory & Cognition*, **32(6)**, 990–1006.

Hayes, B. & Adams, R. (2000). Parallels between clinical reasoning and categorization. In J. Higgs & M. Jones (Eds.) *Clinical Reasoning in the Health Professions*. Edinburgh: Butterworth-Heinemann. pp 45–53.

Howarth, B., Hoben, K., Morris, J., Varley, R., Lee, J. & Cox, R. (2005) Investigating students' clinical reasoning: Analysing think-aloud protocols and integrating multiple-source data. *Proceedings of the 11th Biennial Conference of the European Association for Research in Learning and Instruction (EARLI)*, Cyprus, August. Available at www.tlrp.org/dspace/retrieve/318/Cox+full+paper+conf+04+revised.doc

Klahr, D. (2000) *Exploring science: The cognition and development of discovery processes.* Cambridge, MA: MIT Press.

Klahr, D., Fay, A.L. & Dunbar, K. (1993) Heuristics for scientific experimentation: a developmental study. *Cognitive Psychology*, **24(1)**, 111–146.

Lee, J., Dineen, F., McKendree, J. & Mayes, T. (1999) Learning vicariously in a distributed environment. *Active Learning*, **10**, 4-9.

Lemke, J.L. (1990) *Talking science: Language, learning and values.* Westport, CT: Ablex.

Lum, C. & Cox, R. (1998) PATSy - A distributed multimedia assessment training system. *International Journal of Communication Disorders*, **33**, 170–175.

Lum, C., Cox, R., Kilgour, J., Morris, J. & Tobin, R. (1999) PATSy: A multimedia, web-based resource for aphasiologists and neuropsychologists in research and education. *Aphasiology*, **13**, 573–579.

Mayes, J.T., Dineen, F., McKendree, J. & Lee, J. (2001) Learning from watching others. In C. Steeples & C. Jones (Eds.), *Networked learning: Perspectives and issues*. London: Springer.

McKendree, J, Stenning, K, Mayes, T, Lee J, & Cox, R. (1998) Why observing a dialogue may benefit learning: The vicarious learner. *Journal of Computer-Assisted Learning*, **14(2)**, 110–119.

Ohlsson, S. (1992) Information-processing explanations of insight and related phenomena. In M.T. Keane & K.J. Gilhooly (Eds.), *Advances in the psychology of thinking*, Vol. 1, pp. 1–44. New York: Harvester Wheatsheaf.

Patel, V.L., Groen, G.J. & Fredericson, C.H. [1986] Differences between students and physicians in memory for clinical cases. *Medical Education*, **20**, 3–9.

Skehan, P. (1988) *A cognitive approach to language learning*. Oxford: Oxford University Press.

Suthers, D., Connelly, J., Lesgold, A., Paolucci, M., Erdosne, A., Toth, J. & Weiner, A. (2001) Representational and advisory guidance for students learning scientific enquiry. In K.D. Forbus & P.J. Feltovich (Eds.) *Smart machines in education.* Menlo Park, CA: AAAI Press/The MIT Press.