AI Medical School Tutor: Modelling and Implementation

Shazia Afzal[1], Tejas Indulal Dhamecha[1], Paul Gagnon[2], Akash Nayak[1],
Ayush Shah[1], Jan Carlstedt-Duke[2], Smriti Pathak[3], Sneha Mondal[1],
Akshay Gugnani[1], Nabil Zary[2], and Malolan Chetlur[1]([✉])

[1] IBM Research, New Delhi, India
{shaafzal,tidhamecha,anayak09,snemonda,aksgug22,mchetlur}@in.ibm.com
[2] Lee Kong Chian School of Medicine,
Nanyang Technological University, Singapore, Singapore
{pgagnon,jan.carlstedt-duke,nzary}@ntu.edu.sg
[3] Imperial College, London, UK
smriti.pathak@kcl.ac.uk

Abstract. In this paper we present our experience in the design, modelling, implementation and evaluation of a conversational medical school tutor (MST), employing AI on the cloud. MST combines case-based tutoring with competency based curriculum review, using a natural language interface to enable an adaptive and rich learning experience. It is designed both to engage and tutor medical students through Digital Virtual Patient (DVP) interactions built around clinical reasoning activities and their application of foundational knowledge. DVPs in MST are realistic clinical cases authored by subject matter experts in natural language text. The context of each clinical case is modelled as a set of complex concepts with their associated attributes and synonyms using the UMLS ontology. The MST conversational engine understands the intent of the user's natural language inputs by training Watson Assistant service and drives a meaningful dialogue relevant to the clinical case under investigation. The curriculum content is analysed using NLP techniques and represented as a related and cohesive graph with concepts as its nodes. The runtime application is modelled as a dynamic and adaptive flow between the case and student characteristics. We describe in detail the various challenges encountered in the design and implementation of this intelligent tutor and also present evaluation of the tutor through two field trials with third and fourth year students comprising of 90 medical students.

Keywords: Case-based tutoring · Digital virtual patient · Natural language interface

---

1   Introduction

Virtual case based learning (CBL) exposes students to a variety of clinical scenarios, allowing practice and explicit training in the stages and skills associated with expert clinical reasoning like collecting relevant data, formulating a better abstraction of the patient problem using semantic qualifiers, and the ability to store and recall 'illness scripts' that aid in rapid hypothesis generation and verification leading to a focused comparison of differentials [3]. This implies that the meaningful organization and structuring of knowledge is more important for its contextual retrieval and building of a knowledge-base rather than just the acquisition of foundational science. In fact, the majority of diagnostic errors are known to be caused by cognitive errors that are not related to knowledge deficiency (3%) but to flaws in data collection (14%), data integration (50%) and data verification (33%) [17]. This explains the considerable importance given to clinical exposure early on in medical education. However, traditional modes of imparting such clinical experience through clinical rotations fall short in providing in-depth personalised learning encounters because of coursework schedules and limited availability of expert clinicians. An internal survey conducted by LKC School of Medicine, NTU Singapore found that students often wish to get more time and dedicated one-on-one learning support for clinical cases. Anticipating this gap and thanks to an e-learning grant, LKC School of Medicine, NTU Singapore, partnered with IBM Research to envision an AI driven medical school tutor (MST) based on Digital Virtual Patient (DVP). The aim was to prepare students in the transition from academic study to clinical application in an engaging manner. Real life clinical cases that were culturally and contextually relevant for the target students were selected by subject matter experts (SMEs). Pedagogical strategies for promoting diagnostic skills as recommended in medical literature were chosen to deliver these cases. A conversational interface was deemed appropriate to support interaction in natural language. The clinical cases were authored and annotated in a manner that permitted dialog-based interaction marked with key references to coursework. The medical school curriculum that was organised in the form of learning outcomes (LOs) was linked and mapped to serve as the back-end for driving recommendations in the tutor. Assessment data of the students was retrieved from the school's data warehouse to build mastery profiles of students and seed the learner model. Finally, MST was optimised for mobile device usage in order to maximise value to students who had expressed the need of being able to access it anytime, often on the go. The following sections describe the challenges faced during modelling of different components that eventually led to the implementation and cloud based delivery of MST.

2   Related Work

One of the earliest attempts in the medical tutoring space was the knowledge based system GUIDON [4] for training in infectious diseases. Its rule base was

subsequently reconfigured to create GUIDON 2 that reasoned more like human experts and provided better explanations [5]. Later, other intelligent tutoring systems like MR Tutor [19], SlideTutor [6], CIRCSIM [8], COMET [20], and SIAS [16] were introduced. These catered to very specific topics, but as a proof-of-concept they did report better motivation and engagement in addition to showing significant learning gains. More popular commercial medical education apps like MedScape [9], Prognosis[1] and Human Dx[2] - amongst others, are also worth mentioning because of their wide user base and ease of use. Finally, even though MST has a conceptual resemblance with the clinical reasoning tool described in [10] it differs significantly with respect to being tutor-driven interaction design, use of natural language, resource recommendation, continuous evaluation based just-in-time feedback and other features as described in the following sections. These features we believe can overcome the lower adoption rates and student engagement found in [10].

MST differs from existing tutoring systems on several key aspects. Firstly, it provides a holistic learning experience where medical cases are used to train students in diagnostic reasoning skills in relation to their foundational curriculum. So the focus is not entirely on diagnostic accuracy but rather the diagnostic ability of students to integrate and apply the relevant knowledge from their curriculum in the context of each medical case. Secondly, MST employs a set of diagnostic activities for each case and the assessment of students on each of these is represented in an open learner model (OLM). This assessment is done automatically and dynamically, and can eventually support the tracking of student's diagnostic ability in an evidence-based manner. Thirdly, the underlying knowledge-base built as the foundation of MST has resulted in an entire medical school curriculum being automatically linked and mapped across the entirety of learning outcomes and their associated resources. Finally, each medical case is structured and authored by SMEs to drive an engaging interaction with anchors to suggested reading at strategic points. The outcome is a rich DVP schema that can drive an intelligent conversation with students by replicating a real clinical encounter in terms of case presentation, information flow and knowledge interrogation. The authoring effort is expected to reduce drastically by using semi-automated methods using advanced NLP/AI technology. To sum, MST offers a comprehensive learning experience by leveraging foundational curriculum knowledge with established educational strategies to provide medical students with anytime, anywhere access to relevant medical cases for the ongoing development and refinement of their clinical reasoning skills and competency. The novelty in MST is to build a complete tutoring system that understands the natural language conversations interspersed with clinical terms and connects the background knowledge with curriculum while responding with meaningful interactions to provide a rich learning experience.

---

[1] https://www.medicaljoyworks.com/prognosis-your-diagnosis.
[2] https://www.humandx.org/.
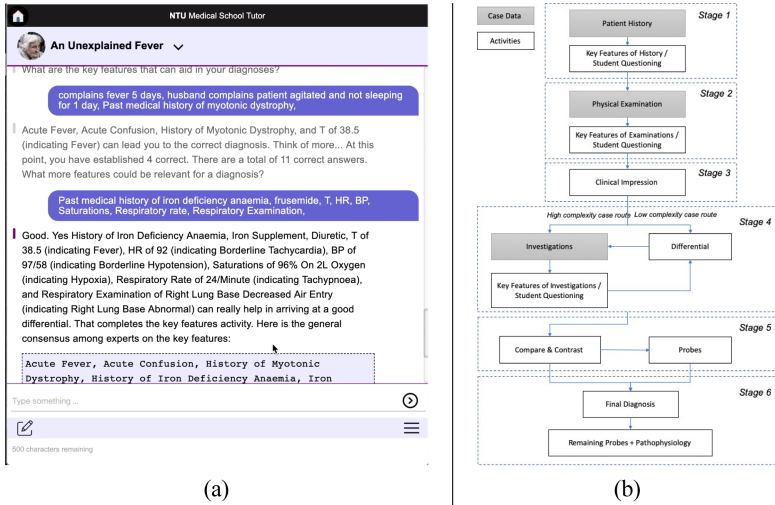
(a)   (b)

Fig. 1. (a) Sample MST interaction for key features activity on a clinical case. (b) Interaction flow indicating sequence of activities for a typical interaction around a DVP.

3   Case Representation and Modelling

CBL is an educational paradigm closely related to Problem based Learning (PBL) in which real-life cases form an authentic context for learning activities meant to promote active problem solving and foster deeper knowledge [13,22]. A DVP is an instantiation of a clinical case that is often used for CBL. DVPs can have varying levels of realism ranging from text based descriptions to high fidelity simulations [12]. While the granularity and design of DVPs depends on their proposed usage, their development and authoring costs remain high making it challenging to scale these for new or unseen cases. MST also has its unique requirement for a DVP: clinical cases need to be authored and annotated in a way that enables a conversational interaction scenario while fulfilling the pedagogical goals of knowledge acquisition, application and reinforcement. The challenge is to not just annotate each case in an elaborate manner but to also ensure that it conforms to the unfolding of a real clinical encounter. The annotation needed to be objectively done so that students performance could be evaluated on each activity. SMEs from LKC School of Medicine carefully selected relevant clinical cases based on real-life patient cases. Together with computer scientists, an annotation protocol was devised and a DVP schema was generated that captured the essential knowledge elements as well as the information flow and related curriculum linkages. This knowledge modelling evolved from custom spreadsheets into DVP JSON (JavaScript Object Notation) objects - a lightweight machine readable data format. This was a non-trivial and time consuming task that involved several rounds of iterations and consultations until a final structure was agreed upon. The resulting DVP schema was a collaborative effort between expert clin-

Table 1. Activities driving interaction with a clinical case.

Activity name	Description
Key features	It trains students in recognizing key clinical findings (e.g. high body temperature) potentially relevant to their differential diagnoses. It aligns with the first stage of data acquisition, mental abstraction and promotes systematic case-reading for distilling pertinent information to formulate possible hypotheses
Clinical impression (CI)	CI is a concise description of patients' key clinical issue(s) with relevant history and examination including functional impairment in clinical terminology (e.g. fever). CI is similar to Problem Representation [21] The purpose is to check students' conceptual understanding of the problem and encourage use of clinical vocabulary that can facilitate retrieval of relevant content from students clinical knowledge/experience
Student questioning	An important part of history taking is the ability to follow a clear line of enquiry and interrogate the patient. To be able to ask the right questions ensures focused data collection which increases the likelihood of potential diagnoses. This activity guides students to ask targeted questions of the patient given the case information presented so far
Knowledge/skills recall	Relating the case data and the path to a possible diagnosis tests medical knowledge and its application. Therefore, probes in the form of multiple answer questions are interspersed in the MST interaction to help students recall and apply the knowledge acquired in the foundational academic years. MST also has provision for open text probes for flexible testing should the SME require
Differential	This involves listing of the three most likely diagnosis in an attempt to identify diseases that may cause patients presentation. The differential diagnosis is the first attempt to distinctly pronounce the potential hypotheses
Compare & contrast	Rather than listing out the features of possible diagnoses, expert clinicians compare and contrast the evidence against possible diagnoses which makes the focus on the discriminating features for each differential. MST uses this method to teach students to differentiate between diagnostic possibilities by reflecting on the relationship among the key clinical features. Contrasting key clinical features against multiple diagnoses leads to prioritizing a final diagnosis and can help in building of appropriate knowledge networks for future application
Final diagnosis	The final diagnosis is the identification of the disease that is most likely the cause of the patients illness. The process of concluding the final diagnosis is as important as arriving at the decision itself so all other activities enable and support this decision making and help students make a conclusion

icians and computer scientists that laid the foundation of the MST. We realised that to make the project scalable and viable in the long run the authoring had to be complemented with proper tooling to make it easy for SMEs to create and review cases. Eventually a complete case authoring tool was developed to semi-automate the process using advanced AI/NLP techniques.

4   Interaction Modeling

MST takes Bowens' [3] elements of the diagnostic reasoning process as the core framework to design and steer learning activities that target student training and practice in: data acquisition skills; problem summarizing using clinical vocabulary; analysing competing hypotheses; prioritizing of diagnostic possibilities; creation of illness scripts by integrating contextual case knowledge; understanding prototypical presentations of diseases; retrieval and recall of acquired curriculum knowledge; encouraging targeted and timely reading; and finally reinforcing the foundational clinical knowledge with clinical practice. The following sections describe the activities that enable development of these skills and how they are interleaved with the DVP to create an interactive seamless flow.

4.1   Learning Activities and Case Flow

Figure 1b illustrates the stages within MST and how the various clinical reasoning activities are surfaced. The stages correspond to blocks or gates that end with learning points and takeaways. This sequence template tries to replicate the stages of a real clinical encounter and was created in consultation with SMEs. Any variations on this case flow are defined by SMEs as part of the authoring process. Table 1 provides a brief description of the specific activities. In order to make the interaction crisp and engaging, students are given only two attempts to specify their answers. Exception to this are the probing questions that are delivered in a multiple choice question answer style of interrogation. The expert answers with explanations are provided at the end of each activity. These help students introspect and validate their understanding.

4.2   Response Generation

MST models conversation with students using the above-mentioned case flow to achieve the tutor goals of case presentation, activity sequencing, activity evaluation, resource recommendation and simultaneously building of a student model. A major challenge is to meet the pedagogical goals while maintaining conversational efficiency and providing constructive remediation to the student. MST uses a state space model design to generate tutor responses depending on the state and location in the dialog. A library of responses collectively compiled by experts are used to seed the tutor responses in order to ensure culturally and professionally appropriate language. A dialog is enabled by alternate turn-taking where student responses are evaluated in the context of ongoing task and

recognised intent. Expectation setting occurs at the beginning of each activity when students are given activity specific instructions. An appropriate response is accordingly generated which consists of an acknowledgment, an evaluation statement and a dialog advancer to keep the conversation going. At the end of each activity the tutor response also includes an evaluation component to show the score attained. Refer to Fig. 1a for an example of how MST responds to student input by acknowledging the response, asserting correct information and also taking the dialog forward. The response generation has scope to include a students performance as well as engagement parameters so that tutor responses are specifically aligned to the actual state of a student.

5   Context and Domain Modeling

Understanding the students' natural language input in the context of a clinical case and its medical vocabulary is one of the main challenges for MST. The conversational agent within MST has to understand the intention of the student and has to respond appropriately to retain engagement. Responses that do not meet student's expectations or are incorrect will result in disengagement. The tutor's response has to be relevant and valuable with respect to the immediate learning activity. Thus, understanding the student's intent is essential for ensuring smooth interaction and superior experience.

In MST, students' input are modelled as a bag of intents. Intents are small group of one-to-four words uttered during interaction with the tutor. There may be any number of intents within a student response. MST relies on the Watson Assistant service [1] to identify an intent given a group of words. MST identifies all the different intents (bag of intents) within a student response and their appropriateness in the ongoing conversation. The bag of intents are returned to tutoring engine for further processing by Input Curator module within MST. This machine learning module relies on Unified Medical Language System [2] knowledge-base for training and understanding medical terms. For example, when a student writes `high temperature` or `fever` or `pyrexia`, it should all be understood as the same intent. Figure 1a shows the correctly expressed intents are also part of system response in addition to unexpressed intents (shown in bold face). This is to further inquire the student to provide missing parts of the answer.

6   Content Modeling and Curriculum Linkage

Appropriate content recommendation customized to each learner is a crucial feature in MST. This is achieved by linking each clinical case with the underlying curriculum content in the form of Learning Resources (LRs) via Learning Objectives (LOs). MST uses advanced NLP techniques to extract all related LOs (LO-LO Relatedness prediction) and identifies pages of relevant LRs (LO-LR Relevance prediction) to relate concepts from the case with those in the curriculum. The detailed modeling, experiments and results of LO-LO relatedness prediction and LO-LR relevance prediction can be found in [15,18] respectively.

6.1   LO-LO Relatedness

Predicting LO-LO relationship requires looking into the semantic content of disparate LOs, in addition to their relatedness in the curriculum hierarchy [15]. This is formulated as a three-class classification task where given a pair of LOs, a classifier is trained to categorize them as being either strongly related, weakly related, or unrelated based on annotations obtained from Subject Matter Experts (SMEs). The representation for a pair of LOs is the concatenation of their curriculum and semantic features, that encode their relative position in the curriculum hierarchy and similarity in meaning, respectively. This feature-set is used with a random forest classifier, that learns to classify the relationship between pairs of LOs. This LO relationship extraction system is then applied to uncover LOs relevant to clinical cases (DVP) in MST so that useful LRs can be recommended.

6.2   LO-LR Relevance Prediction

Identifying relevance of an LR page with an LO has three key aspects–(1) Lexical, (2) Semantic, and (3) Spatial [18]. The overlap between the terms in the LO and those in the page can be used to identify some of the relevant pages. Similarly, semantic overlap between the LO and the page can identify relevant pages that are not identified by lexical matches. Additional pages that are adjacent to highly relevant pages (either lexical or semantic) are also likely to be relevant. Exploiting this spatial aspect identifies additional relevant pages to the existing set of pages. This problem is formulated as a binary classification task, where given an LO and a page of an LR, a machine learning model predicts its relevance. It would be ideal to train a joint model that is able to utilize all three aspects (lexical, semantic, spatial) to make relevance predictions. However, obtaining annotated training data for this task is expensive as the SME has to go through the entire LR for annotating each LO. Thus, a pipelined approach consisting of separate models capturing each of the aspects is used to convert the alignment problem into a page relevance classification problem.

7   Learner Modeling

The student model for MST tracks knowledge and behaviour including estimation of mastery level from prior assessment data, performance on the various reasoning activities, and behaviour based implicit metrics of engagement. The idea was to build a student profile that could enable open learner model visualisation, provide adaptive learning pathways, allow personalised feedback generation, and tailored sequencing and recommendation of content - eventually leading to customising interactions and optimising the tutor strategy overtime.
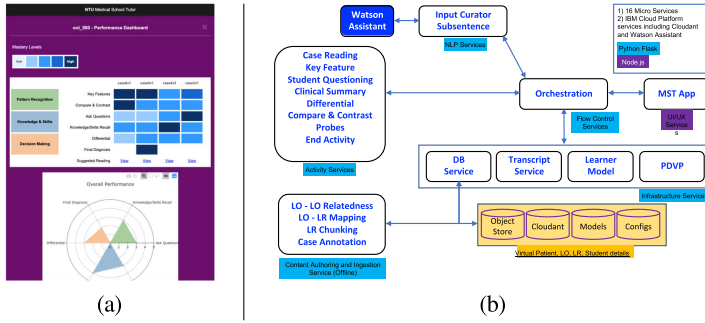
Fig. 2. (a) Learner model visualisation (b) Different components of MST and their interaction.

7.1 Representation

The knowledge and behaviour models in MST have two components to separate the prior student profile from the MST based profile. A combination of overlay and stereotype modelling is used for prior mastery estimation by conducting IRT [11] analysis on the historical performance data from the medical schools' formative assessments called IRA (Individual Readiness Assessment). Students are categorised into five levels of mastery ranging from beginner to expert using overall, cohort-level, year-level, block-level, and LO level IRT estimates. The prior behaviour model is captured in the form of various metrics like login frequency, dwell time, media preferences and resource access to get an idea of the interaction style of students.

MST tracks detailed activity level performance at case level and an overall global estimation across all cases for each student. The activity level assessments form the basis of an evaluation model that bins students into five mastery levels. Activity level scores are computed by comparing student answers against the corresponding expert answers. Additionally, an overall reasoning ability is approximated using dimensions of Pattern Recognition, Knowledge and Skills, and Decision-Making using scores across cases. Key features and Compare & Contrast scores give an estimation of the Pattern Recognition skills of the student. Differential and Final Diagnosis give an indication of the Decision-Making skills while the Probes showcase Knowledge & Skills. Taken together, these three higher order measures highlight the various components of diagnostic reasoning ability and help in identifying areas of strengths and weaknesses. The MST learner model can be instrumental in getting first evidence on automatic assessment of diagnostic reasoning.

7.2 Visualisation

In the spirit of OLMs [7], MST provides students with a powerful visualisation of their student model in order to motivate and track their learning and help

them reflect on their performance Two data visualisation techniques are used in the learner dashboard. The first one is based on the Mastery Grid [14] concept wherein the student mastery levels are depicted using colour gradients in a grid format. The second visualisation is a spider diagram that provides a summary view of higher order skills of pattern recognition, knowledge and skills, and decision-making. The arcs represent the different activities and their gradations correspond to the mastery levels. The scores for this visualization are aggregated across all the cases completed by the student so that it gives an overview of reasoning ability in an intuitive manner. Figure 2a shows how the dashboard looks for a dummy user who has gone through four cases and completed one out of them.
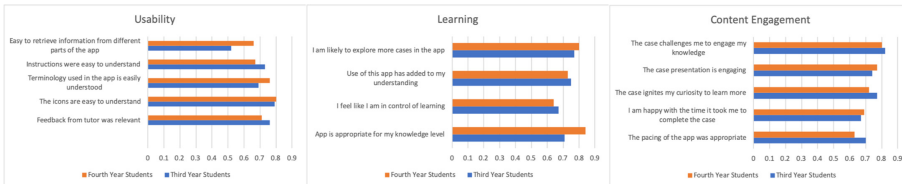

Fig. 3. Feedback survey results

8   Implementation Details

MST is implemented in a cloud-native microservice-based architecture on IBM Cloud. Figure 2b shows its key components and how they interact. The Input Curator service analyses each student input to identify the intent and answer. This service uses advanced NLP techniques and relies on IBM's Watson Assistant service [1]. The identified intents and current state of student in the interaction flow 1b is passed back to the Orchestration which in turn delegates control to the appropriate activity specific microservices. The Learner Model service updates the performance and engagement parameters after completion of individual activities in the interaction flow - refer to Fig. 1b. Additional offline components that complement the tutor system to support case authoring, LO-LO relation extraction, LO-LR mapping, LR chunking and training of Watson Assistant are shown separately in the figure.

9   Experimental Validation

We conducted two field trials with medical students to validate the goals of MST and evaluate its efficacy. The purpose of these trials was not to compare these two user groups but to assess the functioning of the system and get direct feedback on its perceived value from end-users. The trials were conducted a few weeks apart so that feedback taken from students in the first field trial was

incorporated into MST before conducting the second trial. The first field trial was conducted with 55 third year medical students while the second trial was done with 35 fourth year medical students. All students voluntarily participated and were given a brief introduction about the purpose of the experiment. More than 75% of students used laptops while less than 10 % of students used their smartphones during the trials. All students worked on the same dengue case in all sessions. Student's participation was monitored through a muted video conference session to avoid distraction. All students appeared to be considerably engaged while working through the cases in MST. A feedback survey questionnaire was completed by all participants in both trials. It had 14 Likert questions aligned to the dimensions of usability, learning gains and content engagement. Additionally, there were two open style questions asking students about what they liked in MST and what they wished to see in future improvements. The results of the survey comparing feedback between the two groups are depicted in Fig. 3 with scores normalized on 0–1 scale. All scores for usability, learning and content engagement are well above 0.6 with the exception of one score of 0.52 given in the first field trial on ease of finding information on the app. The change of this score to over 0.7 in the second field trial shows that the feedback was taken constructively and utilised to improve the app. Overall, all students rated MST over 0.65 on usability and experience. On the dimension of learning, students gave a score of more than 0.7 vindicating the value of MST as a supplement to learning. Students found themselves in control of their learning, and expressed their interest to explore more cases. Students also found the cases both challenging and engaging. They found the pacing of activities appropriate and were generally happy with the time it took to complete a case. Students comments from the open style questions gave interesting insights into their experience and views on MST. Figure 4a presents the word cloud of the notable features liked by students. For improvements, students mainly expected clarity in instructions, better response classification and an ability to ask counter questions about the systems evaluation or explanation of answers. We aim to follow up on these findings and suggestions prior to conducting a longitudinal study with students. We would especially focus on using the learner model to compare learning gains and changes in diagnostic skills overtime.

Both field trials collectively gave us about 50 h of interaction data. Students spent an average of 33 min on the dengue case with the least time spent being 10 min and the maximum 50 min. Considering time spent as a proxy measure of student engagement, we can perhaps conclude that students were motivated and all of them completed the case till the end. Figure 4b shows the distribution of overall time spent on the case as well as in individual learning activities. Students appear to have spent more time in key features, clinical impression and student questioning activities.

To sum, the survey results and data from these initial evaluations are positive and encouraging. Usability, learning gains and engagement seems to be high among the students from the field trials. Students also proposed areas of improvement in MST specifically regarding instructions not being clear and not
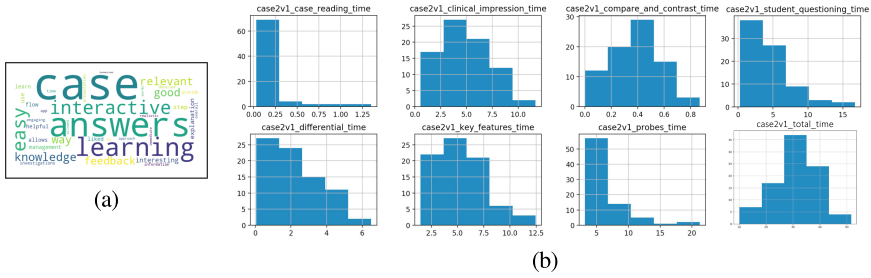
Fig. 4. (a) Word-cloud of what students liked (b) Distribution of time-spent in MST (in mins)

knowing the response expected by the tutor. Students also cited instances where the tutor was not able to understand their responses. Understanding nuances in natural language and medical vocabulary is a huge NLP challenge and this is being currently worked at to improve the experience. However, the reported limitations do not come in the way of achieving a superior learning experience.

10   Summary

We described the Medical School Tutor (MST) as a holistic learning tool driven by AI techniques to prepare medical students in their transition from academic study to clinical application. The interaction design modeled on CBL pedagogy engaged students with realistic clinical cases. The content and context understanding enabled using NLP techniques allows seamless introduction of curriculum resources and generation of appropriate feedback. The survey results show high scores on usability, learning gains and engagement; and validate the purpose and efficacy of conversational tutoring in complex medical domain. The combination of clinical cases with medication foundation curriculum through conversational interface seems to be a novel and an enriching experience for students. In future, MST aims to provide a truly personalized tutoring experience featuring responses adapted to student's learner model.

References

1. IBM Watson Assistant. https://www.ibm.com/cloud/watson-assistant/
2. Bodenreider, O.: The unified medical language system (UMLS): integrating biomedical terminology. Nucleic acids Res. **32**, D267–D270 (2004)
3. Bowen, J.L.: Educational strategies to promote clinical diagnostic reasoning. New Engl. J. Med. **355**(21), 2217–2225 (2006)
4. Clancey, W.J.: Guidon. J. Comput.-Based Instr. (1983)
5. Clancey, W.J., Letsinger, R.: NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching. Stanford University Stanford, Department of Computer Science (1982)

6. Crowley, R.S., Medvedeva, O.: An intelligent tutoring system for visual classification problem solving. Artif. Intell. Med. **36**(1), 85–117 (2006)
7. Desmarais, M.C., Baker, R.S.: A review of recent advances in learner and skill modeling in intelligent learning environments. User Model. User-Adap. Inter. **22**(1–2), 9–38 (2012)
8. Evens, M.W., et al.: Circsim-tutor: an intelligent tutoring system using natural language dialogue. In: Conference on Applied Natural Language Processing: Descriptions of System Demonstrations and Videos (1997)
9. Frishauf, P.: Medscape-the first 5 years. Medscape General Med. **7**(2), 5 (2005)
10. Hege, I., Kononowicz, A.A., Adler, M.: A clinical reasoning tool for virtual patients: design-based research study. JMIR Med. Educ. **3**(2), e21 (2017)
11. Johns, J., Mahadevan, S., Woolf, B.: Estimating student proficiency using an item response theory model. In: International Conference on Intelligent Tutoring Systems, pp. 473–480 (2006)
12. Kononowicz, A.A., Zary, N., Edelbring, S., Corral, J., Hege, I.: Virtual patients-what are we talking about? A framework to classify the meanings of the term in healthcare education. BMC Med. Educ. **15**(1), 11 (2015)
13. Lajoie, S.P., Faremo, S., Wiseman, J.: Tutoring strategies for effective instruction in internal medicine. Int. J. Artif. Intell. Educ. **12**(3), 293–309 (2001)
14. Loboda, T.D., Guerra, J., Hosseini, R., Brusilovsky, P.: Mastery grids: an open source social educational progress visualization. In: European Conference on Technology Enhanced Learning (2014)
15. Mondal, S., et al.: Learning outcomes and their relatedness in a medical curriculum. In: Innovative Use of NLP for Building Educational Applications, pp. 402–411 (2019)
16. Muñoz, D.C., Ortiz, A., González, C., López, D.M., Blobel, B.: Effective e-learning for health professional and medical students: the experience with sias-intelligent tutoring system. Stud. Health Technol. Inform. **156**, 89–102 (2010)
17. Nendaz, M., Perrier, A.: Diagnostic errors and flaws in clinical reasoning: mechanisms and prevention in practice. Swiss Med. Weekly **142**(4344), (2012)
18. Saha, S., Chetlur, M., Dhamecha, T.I.: Aligning learning outcomes to learning resources: a lexico-semantic spatial approach. In: IJCAI (2019)
19. Sharples, M., Du Boulay, J., Teather, B., Teather, D., Jeffery, N., Du Boulay, G.: The MR tutor: computer-based training and professional practice. In: Cognitive Science Research Paper (1995)
20. Suebnukarn, S., Haddawy, P.: Comet: a collaborative tutoring system for medical problem-based learning. IEEE Intell. Syst. **22**(4), 70–77 (2007)
21. Ten Cate, O., Custers, E.J., Durning, S.J.: Principles and Practice of Case-Based Clinical Reasoning Education: A Method for Preclinical Students, vol. 15. Springer, Heidelberg (2017). https://doi.org/10.1007/978-3-319-64828-6
22. Williams, B.: Case based learning—a review of the literature: is there scope for this educational paradigm in prehospital education? Emerg. Med. J. **22**(8), 577–581 (2005)