Learning and Teaching with Conversational Agents

Edith Law
Elizabeth Nilsen
Ken Jen Lee
Apoorva Chauhan
Jessy Ceha
Parastoo Baghaei Ravari
Amy Debbane
Kexin Yang
University of Waterloo
Waterloo, Ontario, Canada
edith.law@uwaterloo.ca
enilsen@uwaterloo.ca
kj24lee@edu.uwaterloo.ca
apoorva.chauhan@uwaterloo.ca
jceha@uwaterloo.ca
parastoo.baghaei.ravari@uwaterloo.ca
agdebbane@edu.uwaterloo.ca
k236yang@uwaterloo.ca

CHI'20,, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
This position paper describes our research and teaching initiatives around conversational agents, and the insights they provide about the grand challenges of conversational agents as a research area.

Author Keywords
conversational agents; learning by teaching; curiosity; question asking; graduate and elementary education

CCS Concepts
•Human-centered computing → User studies; •Applied computing → Interactive learning environments; Collaborative learning;

Introduction
Conversational agents have been used in a variety of ways in education, to handle administrative tasks [7], motivate and support students [5, 10], as well as assess and enhance learning [6, 16, 9, 2, 15, 8, 12, 11]. Such conversational agents can take many forms, as text-based chatbots [15], voice agents [19] or physical robots [17]. They can also adopt different roles in the learning process—as a teacher [1], a peer [3], or a less knowledgeable learner taught by students [14, 13].

This position paper describes (1) our research and development of conversational agents used in education, designed to enhance students' curiosity [3], foster question asking skills [1], and support learning by teaching [13, 4]; (2) a graduate course (CS889) developed at University of Waterloo on conversational agents. We conclude by discussing the grand challenges as informed by these research and teaching activities around conversational agents.

Agent as Peer: Curiosity

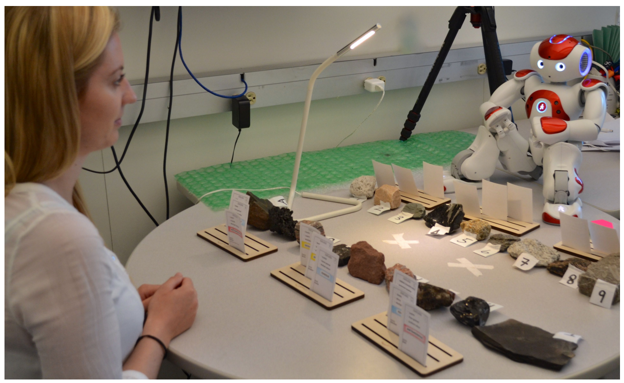

Figure 1: Agent verbally expressing its curiosity about rocks during the LinkIt! game in order to influence the student's curiosity

In [3], we explored the idea of using a conversational robot to enhance students' curiosity about a topic. In particular, we investigated whether a social peer robot's verbal expression of curiosity can affect students' own emotional experience and behavioural expression of curiosity. We designed a rocks and minerals game called LinkIt! (Figure 1), in which the participant and the robot takes turn guessing which rocks have similar features or belong to the same category, and thus can be "linked". In a between-subject experiment with 30 participants, we tested a robot that verbally expresses curiosity, curiosity plus rationale or no curiosity. Results show that curiosity can be 'contagious'—the robot that verbally expresses curiosity was able to influence participants to feel and act curious.

Agent as Teacher: Ability to Ask Questions

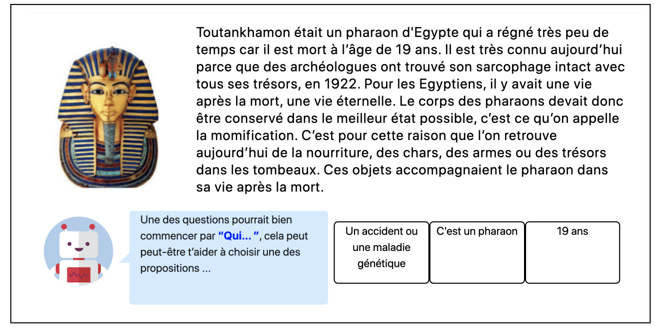

Figure 2: Agent influencing the complexity of the questions that students generate by prompting them to choose an answer that would produce a divergent-thinking question.

In [1], we investigated how an agent can be designed to help children practise question asking, which is an important tool for constructing academic knowledge and a self-reinforcing driver of curiosity. Research has found that children's questions in the classroom are infrequent and often superficial. We developed a pedagogical conversational agent that encourages children to ask more complex (i.e., divergent-thinking) questions. Figure 2 shows how the conversational agent influences the type of questions students generate. In each round, students are presented with an article (e.g., about the Egyptian pharaoh Tutankhamun) and a set of propositions, and are asked to choose one of the propositions (e.g., "19 years old") and generate a valid question from it (e.g., "What age was Tutankhamun when he died?"). The agent can influence students to choose a particular proposition, thereby also the

Figure 4: Students using Curiosity Notebook to teach the conversational robot.

Figure 5: Artifacts for supporting learning by teaching.

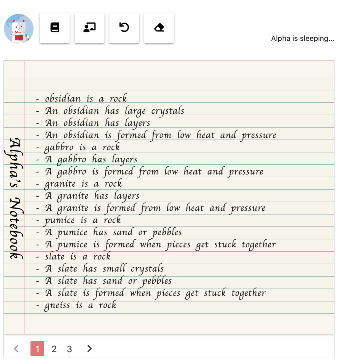

Figure 6: Robot's Notebook

type of questions they would generate, by presenting them with a question-starter prompt, such as "what", "where", "how", etc. We conducted a study with 95 fifth grade students in France, who interacted with an agent that encourages either convergent-thinking or divergent-thinking questions. Results show that both interventions increased the number of divergent-thinking questions and the fluency of question asking. In addition, children's curiosity trait has a mediating effect on question asking under the divergent-thinking agent, suggesting that question-asking interventions must be personalized to each student based on their tendency to be curious.

Agent as Learner: Classifying Objects

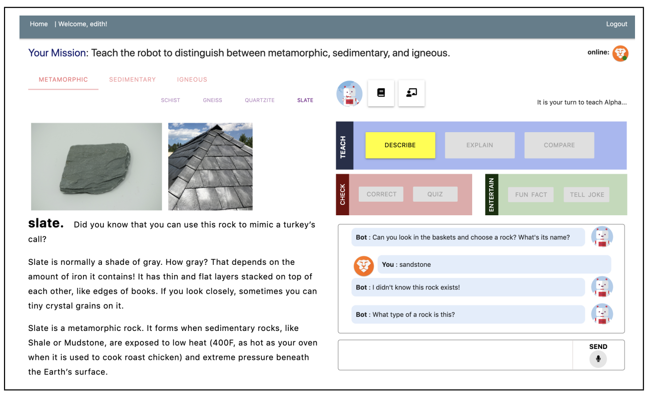

Figure 3: Curiosity Notebook

In the ongoing Teachable Robot Project [13], we are developing a learning-by-teaching platform called the **Curiosity Notebook**, which allows students to work individually or in groups to teach a conversational agent a classification task in a variety of subject topics. Such an agent, who acts as a less knowledgeable peer taught by users, is often referred to as a *teachable agent*. Variants of this platform have been used already in several studies. In [4], we conducted a series of experiments investigating how crowd workers learn to classify text documents by teaching an agent. In [13], we conducted a 4-week exploratory study with 12 fourth and fifth grade elementary school children at a local elementary school, who taught a conversational robot how to classify animals, rocks/minerals and paintings. Our learning-by-teaching platform provides functionalities that allow the conversational agent to be configured as a text-based chatbot, voice-only agent, or physical robot (e.g., NAO), thus enabling researchers to study the learning-by-teaching phenomenon across different types of agent embodiment. Our goal is to deploy this platform to local elementary schools, to study how agent characteristics (e.g., personality, adaptivity) affects the learning-by-teaching experiences and outcomes.

The latest iteration of the Curiosity Notebook, as shown in Figure 3, provides a set of buttons for students to initiate different types of conversations with the agent, which includes—*teaching conversations*, which involve describing the features of an object (e.g., "Schist has layers"), explaining why such features exist (e.g., "the layer comes from piling up of sediments"), and comparing different objects (e.g., explaining the differences between Schist and Shale); *checking* conversations, which allows students to correct the agent on facts that it has learned wrong, or quiz the agent to see how well it can classify the objects; *entertaining* conversations, which asks students to find a fun fact or tell a joke. As the agent is learning, it takes notes and record facts in its notebook (Figure 6).

Graduate Course on Conversational Agents

The graduate course on conversational agents (CS889) developed at University of Waterloo[1] explores recent HCI re-

[1] http://edithlaw.ca/teaching/cs889/w20

search on conversational agents. The course serves simultaneously as a methodology course, where students learn about different study designs (e.g., experiments, diary studies, interviews, etc) and analysis techniques (e.g., statistical modeling, grounded theory analysis), and apply them to research questions related to conversational agents through a substantial course project. The course also involves weekly presentations and critiques of recent HCI papers on conversational agents, to help students understand the scope of this research area. Students have diverse research background, e.g., HCI, machine learning, security and privacy, software engineering; they are asked to tailor their course project to their domain expertise.

Insights about Grand Challenges

Several themes of grand challenges merged from these research and teaching initiatives. One of the biggest challenges is in designing the verbal behaviour of the agent—what the agent says and how—such that it accurately depicts the personality, abilities, attitudes and values that the agent embodies. In the context of the teachable robot project, for example, what would a curious agent learner say? Would the users actually perceive the agent to be curious? How frequently should a humorous agent makes jokes and when? What would be the difference between a self-conscious vs confident agent learners? As the agent learns and grows, how would we script the verbal expressions in such a way that the evolution of the agent learner is believable? When using a physical robot, how do we design the embodiment (e.g., physical form, gestural expressions) to match the role and personality of the agent? The small size and childlike features of the NAO robot, for example, have shown to elicit care-taking behaviour from users, in a way that a voice-only agent may not be able to.

These questions suggest that to anthropomorphize an agent, we need to know how to reproduce human qualities by systematically manipulating the agent's features in such a way that these human qualities will be perceived by the users in the agent as intended. This is challenging because we may not know the exact correspondence between human qualities and agent features that cue the perception of these qualities. As a result, the agent can inadvertently convey something very different from what was intended—for example, in [3], some participants perceived the neutral robot to be curious because it tilted its head, even though this effect was not intended by the designers. One can argue that these are *character design* questions, which academic researchers are ill-equipped to handle, and are better left for authors and screenwriters. This also implies that conversational agent research may require collaboration between computer scientists, behavioural psychologists and creative writers.

There are also methodological challenges for studying conversational agents. First, there is the lack of baselines for comparing conversational agents. What makes one conversational agent better than another? How can we objectively measure the desired qualities of conversational agents? Some of these desired qualities are context-dependent and task-specific. For example, one would want to measure learning performance in the case of an educational conversational agent, but the feeling of productivity in the case of a conversational agents aimed to support work-life transitions [18]. However, other performance measures may be universally relevant, e.g., the consistency of tone in the conversation. It would be interesting to establish and publish a set of common performance measures for conversational agents, with accompanying guidelines on how to systematically assess these measures.

Finally, deeper analysis techniques can be introduced to understand human-agent communication. Beyond task-based outcomes and experience sampling (e.g., keeping counts of certain types of communicative behaviours), one can analyze the moment-to-moment exchanges to characterize other aspects of communication, such as reciprocity and complementarity (e.g., the 'fit' or 'match' between a user and an agent's communicative styles). A related challenge is that users and agents mutually influence each other during a conversation. For instance, if a child asks the robot a lot of questions, then the robot is going to give more information. As result, different children interacting with the same robot might end up seeing different behaviour, not because the robots were programmed to be different, but because the child *pulled* for this difference to emerge. Thus, dyadic models may prove to be more appropriate for studying the interaction between users and conversational agents than simple comparative methods.

REFERENCES

[1] M Alaimi, E Law, K Pantasdo, P-Y Oudeyer, and H Sauzeon. 2020. Pedagogical Agents for Fostering Question-Asking Skills in Children. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3313831.3376776

[2] Vincent Aleven, Kenneth R Koedinger, and Karen Cross. 1999. Tutoring Answer Explanation Fosters Learning with Understanding. In *Proceedings of the 9th International Conference on Artificial Intelligence in Education (AIED '99)*. IOS Press, Amsterdam, 199–206.

[3] Jessy Ceha, Nalin Chhibber, Joslin Goh, Corina McDonald, Pierre-Yves Oudeyer, Dana Kuliundefined, and Edith Law. 2019. Expression of Curiosity in Social Robots: Design, Perception, and Effects on Behaviour. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3290605.3300636

[4] Nalin Chhibber. 2019. *Towards the Learning, Perception and Effectiveness of Teachable Conversational Agents*. Master's thesis. University of Waterloo, Waterloo, Ontario, Canada.

[5] Chih-Yueh Chou, Tak-Wai Chan, and Chi-Jen Lin. 2003. Redefining the Learning Companion: the Past, Present, and Future of Educational Agents. *Computers & Education* 40, 3 (2003), 255–269. DOI: http://dx.doi.org/10.1016/S0360-1315(02)00130-6

[6] Luke Fryer and Rollo Carpenter. 2006. Bots as Language Learning Tools. *Language Learning & Technology* 10, 3 (2006), 8–14. DOI: http://dx.doi.org/10125/44068

[7] Ashok K. Goel and Lalith Polepeddi. 2016. *Jill Watson: A Virtual Teaching Assistant for Online Education*. Technical Report. Georgia Institute of Technology.

[8] Arthur C. Graesser, Kurt VanLehn, Carolyn P. Rose, Pamela W. Jordan, and Derek Harter. 2001. Intelligent Tutoring Systems with Conversational Dialogue. *AI Magazine* 22, 4 (Dec. 2001), 39–51. DOI: http://dx.doi.org/10.1609/aimag.v22i4.1591

[9] M Grigoriadou, G Tsaganou, and Th Cavoura. 2003. Dialogue-based Reflective System for Historical Text Comprehension. In *Proceedings of Workshop on Learner Modelling for Reflection at Artificial Intelligence in Education, International Conference on Artificial Intelligence in Education*. 238–247.

[10] David Griol and Zoraida Callejas. 2013. An Architecture to Develop Multimodal Educative Applications with Chatbots. *International Journal of Advanced Robotic Systems* 10, 175 (2013), 1–15. DOI:http://dx.doi.org/10.5772/55791

[11] N T Heffernan. 2003. Web-based Evaluations Showing Both Cognitive and Motivational Benefits of the Ms. Lindquist Tutor. In *Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies*, Ulrich Hoppe, Felisa Verdejo, and Judy Kay (Eds.). IOS Press, 115–122.

[12] Neil T. Heffernan and Kenneth R. Koedinger. 2002. An Intelligent Tutoring System Incorporating a Model of an Experienced Human Tutor. In *Intelligent Tutoring Systems. Lecture Notes in Computer Science, vol 2363.*, Stefano A. Cerri, Guy Gouardères, and Fàbio Paraguaçu (Eds.). Springer, Berlin, Heidelberg, 596–608. DOI: http://dx.doi.org/10.1007/3-540-47987-2_61

[13] Edith Law, Parastoo Baghaei Ravari, Nalin Chhibber, Dana Kulic, Stephanie Lin, Kevin D. Pantasdo, Jessy Ceha, Sangho Suh, and Nicole Dillen. 2020. Curiosity Notebook: A Platform for Learning by Teaching Conversational Agents. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems (Late Breaking Work) (CHI '20)*. 1–8.

[14] Rod D. Roscoe and Michelene T.H. Chi. 2007. Understanding Tutor Learning: Knowledge-building and Knowledge-telling in Peer Tutors' Explanations and Questions. *Review of Educational Research* 77, 4 (2007), 534–574. DOI: http://dx.doi.org/10.3102/0034654307309920

[15] Sherry Ruan, Liwei Jiang, Justin Xu, Bryce Joe-Kun Tham, Zhengneng Qiu, Yeshuang Zhu, Elizabeth L. Murnane, Emma Brunskill, and James A. Landay. 2019. QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 1–13. DOI:http://dx.doi.org/10.1145/3290605.3300587

[16] Bayan Abu Shawar and Eric Atwell. 2007. Fostering language learner autonomy through adaptive conversation tutors. In *Proceedings of the Corpus Linguistics Conference*. 1–8.

[17] Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning. *Journal of Human-Robot Interaction* 1, 1 (2012), 78–95. DOI: http://dx.doi.org/10.5898/JHRI.1.1.Tanaka

[18] Alex C. Williams, Harmanpreet Kaur, Gloria Mark, Anne Loomis Thompson, Shamsi T. Iqbal, and Jaime Teevan. 2018. Supporting Workplace Detachment and Reattachment with Conversational Intelligence. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article Paper 88, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3173662

[19] Ying Xu and Mark Warschauer. 2019. Young Children's Reading and Learning with Conversational Agents. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems (CHI EA '19)*. ACM, New York, NY, USA, 1–8. DOI: http://dx.doi.org/10.1145/3290607.3299035