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

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

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

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
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

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 Waterloo¹ explores recent HCI re-

¹http://edithlaw.ca/teaching/cs889/w20
search on conversational agents. The course serves simul-
taneously as a methodology course, where students learn
about different study designs (e.g., experiments, diary stud-
ies, interviews, etc) and analysis techniques (e.g., statistical
modeling, grounded theory analysis), and apply them to re-
search questions related to conversational agents through a
substantial course project. The course also involves weekly
presentations and critiques of recent HCI papers on conver-
sational agents, to help students understand the scope of
this research area. Students have diverse research back-
ground, 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 chal-
lenges is in designing the verbal behaviour of the agent—
what the agent says and how—such that it accurately de-
picts 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 cu-
rious? 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 expres-
sions 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 quali-
ties 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 be-
cause 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 ar-
gue that these are character design questions, which aca-
demic 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 con-
versational agents. First, there is the lack of baselines for
comparing conversational agents. What makes one conver-
sational 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 conver-
sational agent, but the feeling of productivity in the case of
a conversational agents aimed to support work-life transi-
tions [18]. However, other performance measures may be
universally relevant, e.g., the consistency of tone in the con-
versation. It would be interesting to establish and publish a
set of common performance measures for conversational
agents, with accompanying guidelines on how to systemati-
cally 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 a 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


