Explanation as Question Answering based on User Guides

Ashok Goel, Vrinda Nandan, Eric Gregori, Sungeun An, and Spencer Rugaber
Design Intelligence Laboratory, School of Interactive Computing, Georgia Institute of Technology
ashok.goel@cc.gatech.edu, vrinda@gatech.edu, egregori3@gatec.edu, sungeun.an@gatech.edu, spencer@cc.gatech.edu

Abstract

Explanation of an AI agent requires knowledge of its design and operation. An open question is how to identify, access and use this design knowledge for generating explanations. Many AI agents used in practice, such as intelligent tutoring systems fielded in educational contexts, typically come with a User Guide that explains what the agent does, how it works and how to use the agent. However, few humans actually read the User Guide in detail. Instead, most users seek answers to their questions on demand. This chapter describes a question-answering agent (AskJill) that uses the User Guide for an AI-based interactive learning environment (VERA) to automatically answer user's questions and thereby explains VERA's domain, functioning, and operation. This chapter also presents a preliminary assessment of AskJill in VERA.

Introduction, Background and Goals

AI research on explanation has a long history that dates at least as far back as the rise of expert systems in the 1960s, e.g., DENDRAL (Lindsay et al. 1993). Mueller et al. (2019) provide a recent and comprehensive review of this research. One of the key ideas to emerge out of this early research was the importance of the explicit representation of knowledge of the design of an AI system (Chandrasekaran & Swartout 1991; Chandrasekaran & Tanner 1989): An explicit representation of the design knowledge of an AI system enables the generation of explanations of the tasks it accomplishes, the domain knowledge it uses, as well as the methods that use the knowledge to achieve the tasks. This raised the question of how this design knowledge can be identified, acquired, represented, stored, accessed, and used for generating explanations. One possible answer was to endow the AI agent with meta-knowledge of its own design (e.g., Goel et al. 1996) and enable the agent to generate explanations through introspection of its meta-knowledge. However, much of AI research on expert systems collapsed by the mid-1990s.

Starting in the 1970s, AI research on explanation also encompassed intelligent tutoring systems (Buchanan 2006). Indeed, in the 1990s, given the collapse of AI research on expert systems, the focus of AI research on explanation shifted to intelligent tutoring systems. Unlike the design stance towards explanations adopted by the research on expert systems, research on tutoring systems took a strongly human-centered perspective. This view emphasized the users and the uses of explanations (e.g., Woolf 2007). For example, Graesser, Baggett & William's (1996) describe question-answering as a basic mechanism of generation of explanations in intelligent tutoring systems, where the answers to the questions meet the requirements and expectations of the human users; Aleven & Koedinger (2002) present explanations of reasoning as a source of new knowledge and learning for the users. However, much of this work perhaps lay a little outside mainstream AI research.

Over the last several years, explanation has again entered mainstream AI research (e.g., Gunning & Aha 2019). This is in part because of advances in machine learning, such as deep learning, that have refocused attention on the need for interpretability and explainability of internal representations and processing in AI agents in general (Gilpin et al. 2018; Rudin 2019). However, explanation of knowledge-based AI systems too is important for reasons of fairness, transparency, accountability, trustworthiness, and human understanding and learning.

In this chapter, the authors take the two ideas from explanations in expert systems and tutoring systems mentioned above as our starting points for generating explanations in knowledge systems: (1) Use of the knowledge of the design of an AI agent as the basis for generating explanations, and (2) human-centered question-answering as the basic mechanism for generation of explanations. They add a third idea to this mix: Given that most practical AI agents, for example almost all intelligent tutoring systems, come with a User Guide that contains knowledge about the domain, design and operation of the agent (Ko et al. 2011), might the User's Guide act as a basis for generating explanations? Note that almost by definition, the User Guide contains information about many types of explanations that users want. For example, a User Guide for an AI agent typically contains information about the domain of the agent, the vocabulary for representing the domain knowledge, the tasks and subtasks the agent accomplishes (what it does), the knowledge and the data the agent uses (its basic components), the methods in the agent that use the knowledge to accomplish its tasks (how the agent accomplishes its tasks), as well as the operation of the agent (how to use the agent). However, few humans actually read the User's Guide in any detail (Rettig 1991; Novick and Ward 2006; Mehlenbacher et. al. 2002).

Instead, most users want answers to their questions on demand, as and when needed. Thus, (3) the authors propose to use the User Guide to generate answers to users' questions.

In this chapter, the authors describe the use of a question-answering agent (called AskJill) for generating explanations about an interactive learning environment (named VERA) based on the latter's User's Guide. AskJill is intended to automatically answer users' questions and thereby explain VERA's domain, functioning, and operation. The authors also present a preliminary formative assessment of AskJill in VERA.

VERA, An Interactive Learning Environment

The VERA project addresses the issues of availability, achievability, and quality of online education. Residential students in higher education have access to physical laboratories, where they conduct experiments and participate in research, thus discovering new knowledge grounded in empirical evidence and connecting it with their prior knowledge. Online learners do not have access to physical laboratories, which impairs the quality of their learning. Thus, the authors developed a Virtual Experimentation Research Assistant (VERA for short) for inquiry-based learning of scientific knowledge (An et al. 2020, 2021): VERA helps learners build conceptual models of complex phenomena, evaluate them through simulation, and revise the models as needed. VERA's capability of evaluating a model by simulation provides formative assessment on the model; its support for the whole cycle of model construction, evaluation, and revision fosters self-regulated learning. Given that residential students too have only limited access to physical laboratories, VERA is also useful for blended learning. VERA is available online (http://vera.cc.gatech.edu) for free and public use.

For the domain of ecology, the authors have integrated VERA with Smithsonian Institution's Encyclopedia of Life that is available as an open-source library and software (EOL; Parr et al.). EOL's TraitBank supports ecological modeling in VERA in several ways: it provides (i) the ontology of conceptual relations for conceptual modeling, (ii) knowledge of specific relations among biological species in a given ecological system, and (iii) the parameters for setting up the simulations. Thus in VERA, biological species are modeled using data directly retrieved from EOL such as lifespan, body mass, offspring count, reproductive maturity, etc. Given that the space of simulation parameters can be very large, and a learner may not know the "right" values for the parameters, once the learner sets up the conceptual model using the EOL digital library, VERA further uses EOL's knowledge of biological species to directly set initial values of the simulation parameters. The learner may then tweak the parameter values and experiment with them.

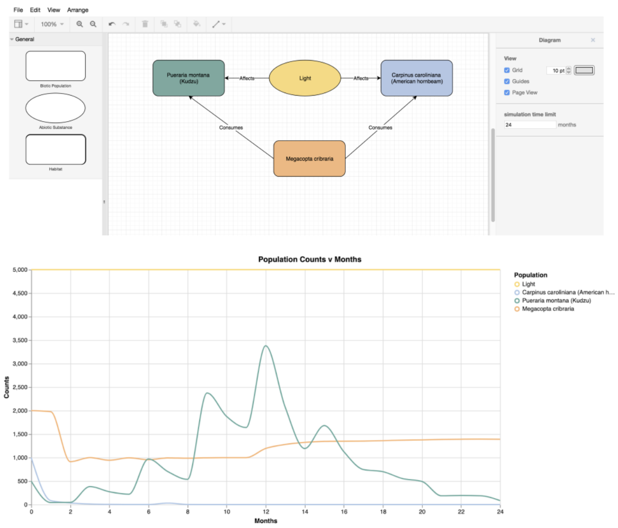


Figure 1. (a) An example of a conceptual model (the top half of the figure) and (b) its agent-based simulation automatically generated by VERA (the bottom half).

Figure 1 illustrates the use of VERA to model the impact of a kudzu "bug" to moderate the impact of kudzu, an Asian invasive species, on the American hornbeam, a kind of tree common in the eastern half of the United States. In Figure 1(a), the learner interactively builds a conceptual model, and in Figure 1(b) VERA illustrates the results of an agent-based simulation of the model. In this case, the simulation results show that because of the introduction of the kudzu bug, the population of kudzu will decline over time and the American hornbeam will survive.

VERA uses agent-based simulations to provide formative assessment on the conceptual models. Note that VERA automatically spawns agent-based simulations from conceptual models: An AI compiler inside VERA understands enough of the syntax and semantics of both the conceptual models and agent-based simulations that it can automatically spawn the latter from the former. This is another example of learning assistance in VERA. This learning assistance enables both student scientists and citizen scientists to model complex phenomena without requiring expertise in the mathematics or mechanics of agent-based simulations. Further, VERA's support for the whole cycle of model construction, evaluation, and revision fosters self-regulated learning.

In 2019, Smithsonian Institution started providing access to VERA directly through the main page on its EOL website (www.eol.org). This means that the hundreds of thousands of EOL users across the world each year, including learners and teachers as well as citizen and professional scientists now have direct access to VERA. This also makes explanations of VERA's domain, functioning and operation critically important.

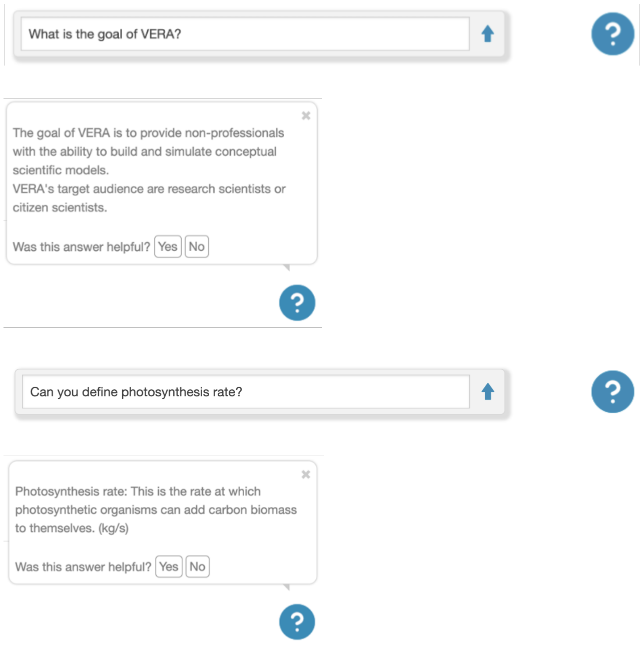
User Guide in VERA

VERA's User Guide and its table of contents is available on its website under the Help section. It includes a written guide describing how users can build and simulate ecological experiments on VERA, the tool's expected behavior, explanations for the vocabulary terms and parameters users can manipulate, and screenshots showing the tool's structure (screens and buttons). Specifically, the 27-page User Guide covers an introduction to VERA, system requirements, steps to access the tool, general approach to build and evaluate a conceptual model of an ecological system, how to use the VERA tool for modeling and simulation (including steps to create a project describing a phenomena and associated models to test various hypotheses), how to use the model editor to manage constituent components and their relationships, how to simulate a model, how to edit model parameters to manipulate results, and ways to get help on the tool.

The User Guide provides illustrative descriptions of the user's workflow on VERA. For example, in its "Getting to know the model editor" section, the User Guide provides an example of a "starter" conceptual model of a simple ecosystem composed of wolves, sheep, and grass, to walk the user through the steps needed to create a the "biotic population" components for each of the three populations. It also shows the user how to define the ecological relationships (destroys, produces, consumes, becomes, affects, can migrate to) between each set of components (e.g. wolves "consume" sheep, sheep "consume" grass), and simulate the model. The User Guide describes how users can set up, start, stop, reset the simulation and export resulting graphs. The User Guide also provides example parameter values showing how parameters can be initialized (Smithsonian's EOL supplies default values) and tuned (provides tuning values) to get the desired population behavior (shows resulting graphs for reference) in the simulation. Last but not the least, the User Guide provides definitions and explanations for commonly used model components (e.g. biotic substance, abiotic substance, and habitat) and their associated simulation parameters (e.g. some parameters for a biotic substance are lifespan, carbon biomass, minimum population, etc.).

AskJill, A Question-Answering Agent

AskJill is a question-answering agent embedded in the VERA interactive learning environment that automatically answers users' questions and thereby explains VERA's domain, functioning, and operation. When a user first logs-in on the VERA website, AskJill welcomes them and prompts them to ask their questions about VERA. The user can type their questions into the AskJill question-answering interface (integrated into the VERA website). AskJill provides accurate answers to the questions within the scope of the User Guide within a few seconds. Figure 2 shows examples of question-answering in AskJill.



What is the goal of VERA?

The goal of VERA is to provide non-professionals with the ability to build and simulate conceptual scientific models.
VERA's target audience are research scientists or citizen scientists.

Was this answer helpful? Yes No

Can you define photosynthesis rate?

Photosynthesis rate: This is the rate at which photosynthetic organisms can add carbon biomass to themselves. (kg/s)

Was this answer helpful? Yes No

Figure 2: A couple of user questions to AskJill about VERA and AskJill's answers to the questions.

AskJill's Generation of an Answer to a Question

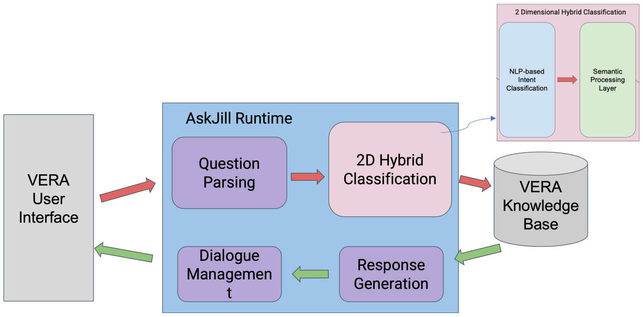


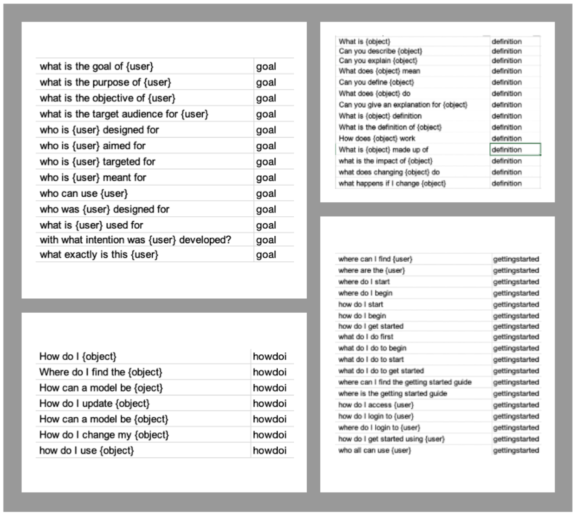
Figure 3: AskJill question-answering data flow diagram

Figure 3 shows AskJill's question-answering data flow diagram. After a user asks a question in VERA's AskJill interface, it is sent to the AskJill system via a REST API. Inside AskJill, the question is parsed, and then sent to a 2D hybrid classification system. The system contains a 2-stage classification process (Goel, 2020). The first is a pre-trained NLP-based intent classification layer that classifies each new question into one of the existing question categories based on user intents. The second is a semantic processing stage that uses the intent to select a rule-based query template. From the 2D hybrid classification system, a query is sent to the VERA's design knowledge database and a response is generated. The response generation system retrieves the associated query response and returns an answer if its confidence value exceeds the minimum threshold (97%). Finally, the dialogue management system post-processes the resulting response, converts it into a “human-like” natural language answer, and sends it back to AskJill in the VERA user interface. After answering, AskJill prompts the user to provide feedback, asking “Was this answer helpful”, and stores the user feedback in her database. That feedback is subsequently used for retraining the agent. If AskJill is unable to answer a question, it can (a) gently redirect the conversation into its domain of competence by suggesting alternate topics associated with the questions it is trained on and/or (b) share relevant links to the User Guide.

Agent Smith: Building AskJill for VERA’s User Guide

AskJill evolved from our earlier work on the Jill Watson project (Goel & Polepeddi 2018) that automatically answered students’ questions on discussion forums of online and hybrid classes. Agent Smith is an interactive generator for generating Jill Watson teaching assistants for different classes (Goel 2020; Goel, Sikka & Gregori 2021): it combines knowledge-based AI, supervised machine learning, and human-in-the loop machine teaching for training a Jill Watson assistant for a new class. Since AskJill for VERA’s User Guide has the same architecture and algorithms as the original Jill Watson for class syllabi, the authors were able to reuse the Agent Smith generator to build the AskJill for VERA. Similar to previous Jill Watson applications, Smith builds a semantic memory for VERA’s vocabulary, system requirements, structure, and tool behavior. It also generates a knowledge base consisting of user intents, keywords, and associated answers. Agent Smith then uses supervised learning to train a classifier to generate an AskJill for VERA. Reusing the Agent Smith technology allows us to train, retrain and generate AskJill agents based on VERA’s User Guide efficiently and easily. AskJill for VERA is encoded in the form of unique question templates related to goals, getting started, definitions, and how-to pointers, simulation parameter default values, and units.

While the rest of the technology from Jill Watson TA (teaching assistant) is reused, Agent Smith utilizes a brand-new set of template questions as well as VERA design knowledge base. A new set of template questions is needed because users pose different related questions (and underlying intents) to AskJill in VERA as compared to course related questions in Jill Watson TA. Similarly, a new knowledge base is needed because the AskJill agent for VERA is based on the User Guide, while the Jill Watson agent is based on course syllabus and schedule. Figure 4 shows an example of the question templates used for training AskJill in VERA. Agent Smith projects the templates onto the VERA ontology and generates the training dataset. The AskJill agent uses the trained model for run-time question answering. Over time, as the authors collect user feedback and analyze missed questions, they can expand the training dataset and retrain AskJill enabling it to answer more and more questions. As a by-product of developing, testing, and training the AskJill Q&A agent, the authors identified definitions and parameters that were initially missing in the User Guide. They have since updated the User Guide to include those missed aspects.



what is the goal of {user}	goal
what is the purpose of {user}	goal
what is the objective of {user}	goal
what is the target audience for {user}	goal
who is {user} designed for	goal
who is {user} aimed for	goal
who is {user} targeted for	goal
who is {user} meant for	goal
who can use {user}	goal
who was {user} designed for	goal
what is {user} used for	goal
with what intention was {user} developed?	goal
what exactly is this {user}	goal

How do I {object}	howdoi
Where do I find the {object}	howdoi
How can a model be {oject}	howdoi
How do I update {object}	howdoi
How can a model be {object}	howdoi
How do I change my {object}	howdoi
how do I use {object}	howdoi

What is {object}	definition
Can you describe {object}	definition
Can you explain {object}	definition
What does {object} mean	definition
Can you define {object}	definition
What does {object} do	definition
Can you give an explanation for {object}	definition
What is {object} definition	definition
What is the definition of {object}	definition
How does {object} work	definition
What is {object} made up of	definition
what is the impact of {object}	definition
what does changing {object} do	definition
what happens if I change {object}	definition

where can I find {user}	gettingstarted
where are the {user}	gettingstarted
where do I start	gettingstarted
where do I begin	gettingstarted
how do I start	gettingstarted
how do I begin	gettingstarted
how do I get started	gettingstarted
what do I do first	gettingstarted
what do I do to begin	gettingstarted
what do I do to start	gettingstarted
what do I do to get started	gettingstarted
where can I find the getting started guide	gettingstarted
where is the getting started guide	gettingstarted
how do I access {user}	gettingstarted
how do I login to {user}	gettingstarted
where do I login to {user}	gettingstarted
how do I get started using {user}	gettingstarted
who all can use {user}	gettingstarted

Figure 4: Some examples of Agent Smith Question Templates for VERA Q&A AskJill Agent

Evaluation of AskJill in VERA

The authors collected AskJill user data both during its use in an ‘Introduction to Biology’ class at Georgia Tech, as well as from citizen scientists discovering VERA through Smithsonian’s website or while browsing the Internet (An et al. 2020, 2021). Currently, AskJill can answer questions belonging to seven categories (intents) of questions shown in Figure 5.

Figure 6 shows examples of a human-generated question from each question category above and as well as AskJill’s responses. The current training dataset consists of 3053 questions both the actual user questions, and anticipated questions from the User Guide.

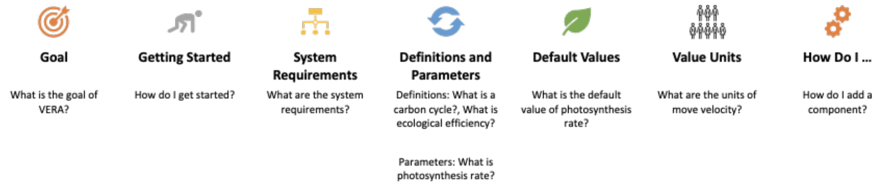
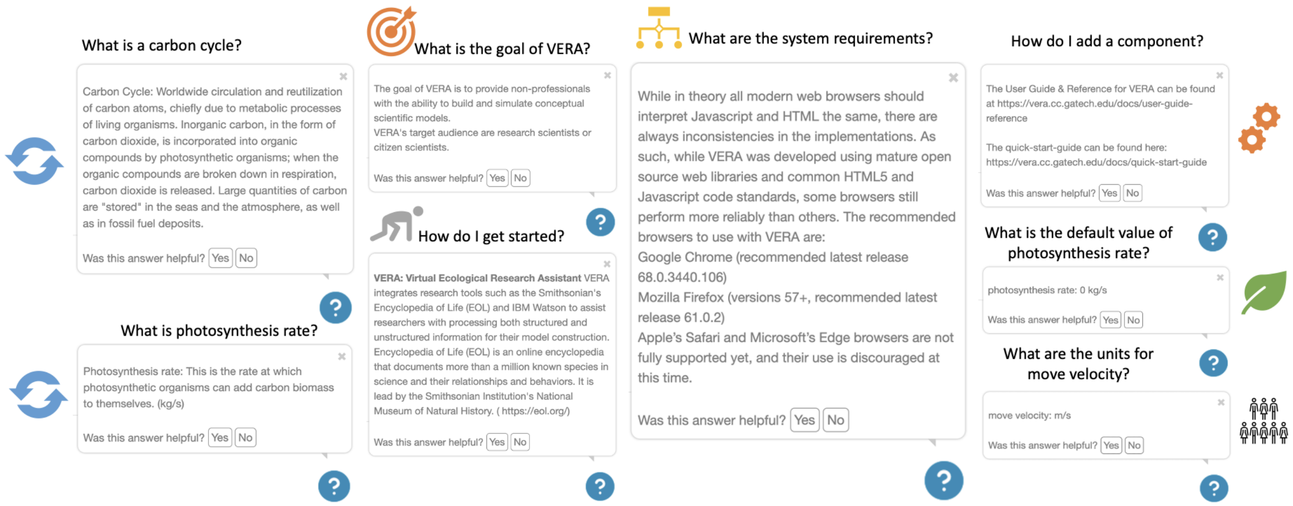

Goal
What is the goal of VERA?
Getting Started
How do I get started?
System Requirements
What are the system requirements?
Definitions and Parameters
Definitions: What is a carbon cycle?, What is ecological efficiency?
Parameters: What is photosynthesis rate?
Default Values
What is the default value of photosynthesis rate?
Value Units
What are the units of move velocity?
How Do I ...
How do I add a component?

Figure 5: User Intent (question) categories on AskJill



What is a carbon cycle?

Carbon Cycle: Worldwide circulation and reutilization of carbon atoms, chiefly due to metabolic processes of living organisms. Inorganic carbon, in the form of carbon dioxide, is incorporated into organic compounds by photosynthetic organisms; when the organic compounds are broken down in respiration, carbon dioxide is released. Large quantities of carbon are "stored" in the seas and the atmosphere, as well as in fossil fuel deposits.

Was this answer helpful? Yes No

What is photosynthesis rate?

Photosynthesis rate: This is the rate at which photosynthetic organisms can add carbon biomass to themselves. (kg/s)

Was this answer helpful? Yes No

What is the goal of VERA?

The goal of VERA is to provide non-professionals with the ability to build and simulate conceptual scientific models.
VERA's target audience are research scientists or citizen scientists.

Was this answer helpful? Yes No

How do I get started?

VERA: Virtual Ecological Research Assistant VERA integrates research tools such as the Smithsonian's Encyclopedia of Life (EOL) and IBM Watson to assist researchers with processing both structured and unstructured information for their model construction. Encyclopedia of Life (EOL) is an online encyclopedia that documents more than a million known species in science and their relationships and behaviors. It is lead by the Smithsonian Institution's National Museum of Natural History. (https://eol.org/)

Was this answer helpful? Yes No

What are the system requirements?

While in theory all modern web browsers should interpret Javascript and HTML the same, there are always inconsistencies in the implementations. As such, while VERA was developed using mature open source web libraries and common HTML5 and Javascript code standards, some browsers still perform more reliably than others. The recommended browsers to use with VERA are:
Google Chrome (recommended latest release 68.0.3440.106)
Mozilla Firefox (versions 57+, recommended latest release 61.0.2)
Apple's Safari and Microsoft's Edge browsers are not fully supported yet, and their use is discouraged at this time.

Was this answer helpful? Yes No

How do I add a component?

The User Guide & Reference for VERA can be found at https://vera.cc.gatech.edu/docs/user-guide-reference

The quick-start-guide can be found here: https://vera.cc.gatech.edu/docs/quick-start-guide

Was this answer helpful? Yes No

What is the default value of photosynthesis rate?

photosynthesis rate: 0 kg/s

Was this answer helpful? Yes No

What are the units for move velocity?

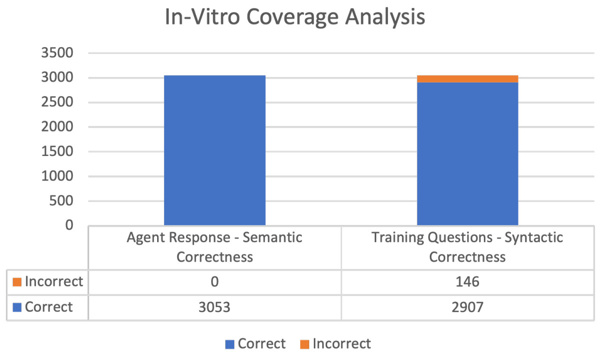
move velocity: m/s

Was this answer helpful? Yes No

Figure 6: Human generated questions and AskJill's agent generated answers.

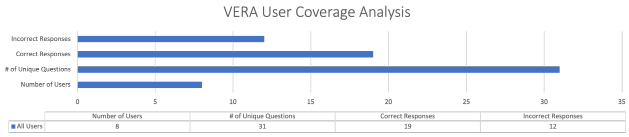
Given that Agent Smith automatically generated the training dataset using a combination of template questions and relevant keywords, the authors also tested for the grammatical correctness of the generated training dataset. Out of 3053 questions, 2907 or 95.2% were accurate. The remaining 4.8% of questions were not grammatically correct but AskJill was still able to resolve the associated intents and answer them correctly. Figure 7 shows our validation results for the current training question set (3053 questions): 100% of the agent generated responses are semantic correct. It also shows the split between syntactically correct and incorrect agent generated questions.

The authors have also collected a small dataset consisting of *in-situ* observations. Figure 8 shows a comparison of data collected from 8 users including external users as well as members of our research laboratory. AskJill correctly answered 19 out of 31 unique questions for all users. They measured user satisfaction using the integrated feedback prompt (Was this answer helpful?) built into the agent's interface and validated that the users confirmed (in some cases there was no feedback) that the correctly answered responses were indeed helpful to the user. Out of the 12 questions that were not answered correctly, a majority are related to simulation parameters, simulation properties, and how-to information specific to a given simulation and thus were outside the competence of AskJill (only 1 out of 12 questions is related to a missed definition). Taking the user feedback a step further, the authors also revised the VERA User Guide to include answers to previously unanswered questions. The closed loop process i.e. adding the information related to missed questions to the VERA knowledge domain, updating the User Guide and retraining AskJill to expand its question answering abilities has resulted in significant improvements to the entire VERA and AskJill pipeline.



	Agent Response - Semantic Correctness	Training Questions - Syntactic Correctness
■ Incorrect	0	146
■ Correct	3053	2907

Figure 7: Agent Response Semantic Correctness and Training Question Syntactic Correctness



	Number of Users	# of Unique Questions	Correct Responses	Incorrect Responses
■ All Users	8	31	19	12

Figure 8: The bar plots show the (a) correct vs incorrect responses (includes "I do not know") responses (b) Number of unique user questions (c) Total number of users.

Discussion

As Mueller et al. (2019) observe, explanations can be of multiple types. Tanner, Keuneke & Chandrasekaran (1993) specifically distinguish between explanations of a phenomenon in the world and self-explanations about an agent's own design. The VERA interactive learning environment, for example, helps users generate explanations of ecological phenomena such as the effect of kudzu bug on the growth of kudzu in the southeast USA; in contrast, AskJill, the question-answering agent embedded in VERA, generates explanations about VERA's domain, design, and operation.

Generation of explanations of an AI agent typically requires specification and encoding of knowledge of the agent's design (Chandrasekaran & Swartout 1991; Chandrasekaran & Tanner 1989). In contrast, AskJill generates answers to a user's questions about an AI agent based on its User Guide, which, for fielded AI agents comes for "free". To put it another way, the authors recast explanation of for practical AI agents as an interactive User Guide for answering users' questions. A corollary here is that the authors seek to identify the design knowledge a User Guide must contain to act as a basis for generating explanations.

While searching the User Guide for the specific information can be laborious and tedious, each information source has its own tradeoffs. On one hand, the AskJill agent provides just in time, curated and accurate answers to the user's questions. On the other, the authors expect the User Guide to offer its readers an opportunity to ponder and deepen their understanding as they search for some specific information and inadvertently discover new knowledge (including context and motivation) due to the inherent differences in the User Guide's structure and format (system diagrams, relationship tables, UI screenshots, related content, and references).

While our approach enables general-purpose explanations, it does not afford explanations of specific instances of reasoning and action by the AI agent. Thus, this approach likely has to be complemented with an episodic approach that relies on specific cases of decision making. Indeed the case-based reasoning research community has developed several schemes for case-based explanation of decision making (Leake & McSherry 2005). In our work along these lines, the authors used meta-cases to capture derivational traces in an earlier interactive learning environment and used the meta-cases to explain the agent's decision making (Goel & Murdock 1996). A future version of AskJill may similarly keep a derivational trace of VERA's decision making and augment its explanatory capability based on a replay of the derivational trace.

Nevertheless, even in its current form, our approach provides insight into specific episodes of decision making both by explaining the vocabulary and the general mechanism of decision making. Consider again the explanation of decisions about the values of the simulation parameters in a specific episode of VERA's agent-based simulation. While AskJill cannot explain why the parameter values led to the specific simulation results in the given episode, it can and does explain each simulation parameter, the role it plays in the simulation, as well as the general mechanism of the agent-based simulation.

As mentioned earlier, AskJill builds our earlier work on the Jill Watson project (Goel & Polepeddi 2018) that automatically answers students' questions on discussion forums of online and hybrid classes. One of the main reasons for the success of Jill Watson is that it took a very human-centric approach: it was trained to answer questions that students had actually asked in online discussion forums over a few years. However, Jill Watson answered questions based on course materials such as class syllabi and schedule, by answering questions based on VERA's Users Guide, AskJill generalizes the approach.

Summary and Conclusions

Explanation of an AI agent requires knowledge of its domain, design and operation. Acquiring, representing, accessing and using this design knowledge for generating explanations is challenging. However, almost all practical AI products and services come with a User's Guide that explains both how the product works and how to use the product. This is especially true for AI agents that actually get fielded in real settings and used by real users. Thus, the authors described the design of a question-answering agent (AskJill) that relies on the User Guide to an interactive learning environment (VERA) to explain its domain, functioning and operation. This means that general explanations of the design of an AI agent now can be generated for "free", without requiring any special encoding of knowledge of the agent's design.

Acknowledgements: Research on the VERA project is funded by an US NSF BigData Grant #1636848 and US NSF AI Institutes Grant #2112532. Both Jill Watson and AskJill use IBM's Watson platform for intent classification; the authors thank IBM for its support for our work. The authors also thank our collaborators on the VERA project including Dr. Jennifer Hammock at Smithsonian Institution and Dr. Emily Weigel at Georgia Tech.

References

Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science,* 26(2), 147–179.

An, S., Bates, R., Hammock, J., Rugaber, S., Weigel, E. & Goel, A. (2020) Scientific Modeling Using Large Scale Knowledge. *In Procs. Twentyfirst International Conference on AI in Education (AIED'2020), pp. 20-24.*

An, S., Broniec, W., Rugaber, S., Weigel, E., Hammock, J., & Goel, A. (2021) Recognizing Novice Learner's Modeling

Behaviors. In *Procs. 19th International Conference on Intelligent Tutoring Systems*. Springer.

Buchanan, B. (2006). A (Very) Brief History of Artificial Intelligence. *AI Magazine* 26(4). pp.53-60.

Chandrasekaran, B., & Swartout, W. (1991) Explanations in Knowledge Systems: The Role of Explicit Representation of Design Knowledge. *IEEE Intelligent Systems*.

Chandrasekaran, B., & Tanner, M. (1989) Explaining Control Strategies in Problem Solving. *IEEE Intelligent Systems.* 4:9-15.

Gilpin, L, Bau, D., Yuan, B., Bajwa, A., Spector, M., & Kagai, L. (2018) Explaining Explanations: An Overview of Interpretability of Machine Learning. In Procs. IEEE Conference on Data Science and Advanced Analytics.

Goel, A. (2020) AI-Powered Learning: Making Education Accessible, Affordable, and Achievable. CoRR abs/2006.01908 (2020)

Goel, A., Gomes, A., Grue, N., Murdock, W., Recker, M., & Govindaraj, T. (1996) Explanatory Interfaces in Interactive Design Environments. In *Procs. Fourth International Conference on AI in Design*, pp. 1-20.

Goel, A., & Murdock, W. (1996) Meta-Cases: Explaining Case-Based Reasoning. In *Procs. Third European Conference on Case-Based Reasoning. Published as Lecture Notes in Computer Science - 1168*, pp. 150-163.

Goel, A., & Polepeddi, L. (2018). Jill Watson. In Dede, C., Richards, J, & Saxberg, B (Eds.) *Learning Engineering for Online Education: Theoretical Contexts and Design-Based Examples*. Routledge.

Goel, A., Sikka, H., & Gregori, E. (2021) Agent Smith: Teaching Question Answering. *In Procs. AAAI 2022 Spring Symposium on Machine Learning and Knowledge Engineering,* Stanford University, March 2022.

Graesser, A., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology,* 10(7), 17–31.

Gunning, D., Aha, D., (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine*, 40(2), 44-58.

Ko, A., Robin., A., et al. (2011) The State of the Art in End-User Software Engineering. *ACM Computing Surveys*. 43 (3): 1–44.

Leake, D., & McSherry, D. (2005) Introduction to the special issue on explanation in case-based reasoning. *Artificial Intelligence Review* 24(2), pp103–108.

Lindsay, R., Buchanan, B., Feigenbaum, E., and Lederberg, J. 1993) DENDRAL: A Case Study in the First Expert System for Scientific Hypothesis Formation. *Artificial Intelligence* 61, 2: 209-261.

Mehlenbacher, B., Wogalter, M., & Laughery, K. (2002). On the reading of product owner's manuals: Perceptions and product complexity. In *Procs. Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, No. 6, pp. 730-734). Sage CA: Los Angeles, CA: SAGE Publications.

Mueller, S., Hoffman, R., Clancey, W., Emery, A., & Klein, G. (2019) Explanation in Human-AI Systems: A Literature Meta-Review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI. https://arxiv.org/abs/1902.01876

Novick, D., & Ward, K. (2006). Why don't people read the manual?. In *Procs.24th annual ACM international conference on Design of communication* (pp. 11-18).

Parr C., Wilson, N., Schulz, K., Leary, P., Hammock, J., Rice, J, Corrigan Jr., R. TraitBank: Practical semantics for organism attribute data. Semantic Web – Interoperability, Usability, Applications 650-1860.

Rettig, M. (1991). Nobody reads documentation. *Communications of the ACM*, 34(7), 19-24.

Rudin, C. (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* 1:206-215.

Tanner M., Keuneke A., & Chandrasekaran B. (1993) Explanation Using Task Structure and Domain Functional Models. In David J, Krivine J., & Simmons R. (eds) *Second Generation Expert Systems.* Springer, Berlin.

Woolf, B. (2007). *Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning.* Morgan Kaufmann Publishers Inc.