



Conversational AIS as the Cornerstone of Hybrid Tutors

Andrew J. Hampton^(✉) and Lijia Wang

University of Memphis, Memphis, TN 38152, USA
andrew.hampton@memphis.edu

Abstract. This paper describes the benefits of artificially intelligent conversational exchanges as they apply to multi-level adaptivity in learning technology. Adaptive instructional systems (AISs) encompass a great breadth of pedagogical techniques and approaches, often targeting the same domain. This suggests the utility of combining individual systems that share concepts and content but not form or presentation. Integration of multiple approaches within a unified system presents unique opportunities and accompanying challenges, notably, the need for a new level of adaptivity. Conventional AISs may adapt to learners within problems or between problems, but the hybrid system requires recommendations at the level of constituent systems as well. I describe the creation of a hybrid tutor, called *ElectronixTutor*, with a conversational AIS as its cornerstone learning resource. Conversational exchanges, when properly constructed and delivered, offer substantial diagnostic power by probing depth, breadth, and fluency of learner understanding, while mapping explicitly onto knowledge components that standardize learner modeling across resources. Open-ended interactions can also reveal psychological characteristics that have bearing on learning, such as verbal fluency and grit.

Keywords: Intelligent recommender · Knowledge components · Triologue

1 Introduction

Adaptive instructional systems (AISs) guide learning experiences in computer environments, adjusting instruction depending on individual learner differences (goals, needs, and preferences) filtered through the context of the domain [1]. These leverage principles of learning to help learners master knowledge and skills [2]. Typically, this type of instruction focuses on one student at a time to be sensitive to individual differences relevant to the topic at hand or instruction generally. It is also possible to have an automated tutor or mentor interact with small teams of learners in collaborative learning and problem solving environments.

Adaptivity in learning technologies dramatically improve upon the paradigm of conventional computer-based training (CBT). Often instruction via CBT operates by simple heuristic that a correct response on a given question calls for a more difficult question to follow, and the inverse. Though this and related techniques can reduce assessment time [3], the adaptivity remains coarse, with no more than basic learning principles applied. In this paradigm, learners may read static material, take an adaptive

assessment, receive feedback on performance, and repeat the process until reaching a threshold of performance. AISs, on the other hand, use fine-grained adaptivity, often within individual problems (step-adaptive), and structure interactions based on empirically-validated pedagogical techniques. Subsequent learning material is typically recommended based upon holistic assessment of mastery or other individual variables (macro-adaptive), rather than the immediately preceding interaction alone. Adaptivity both within and between problems forms two distinct “loops” of adaptivity [4] that can be used in conjunction for more intelligent interactions (see Fig. 1). AISs can also track detailed learner characteristics (e.g., knowledge, skills, and other psychological attributes) and leverage artificial intelligence, informed by cognitive science, to create computational models describing learners and recommending next steps [4–6].

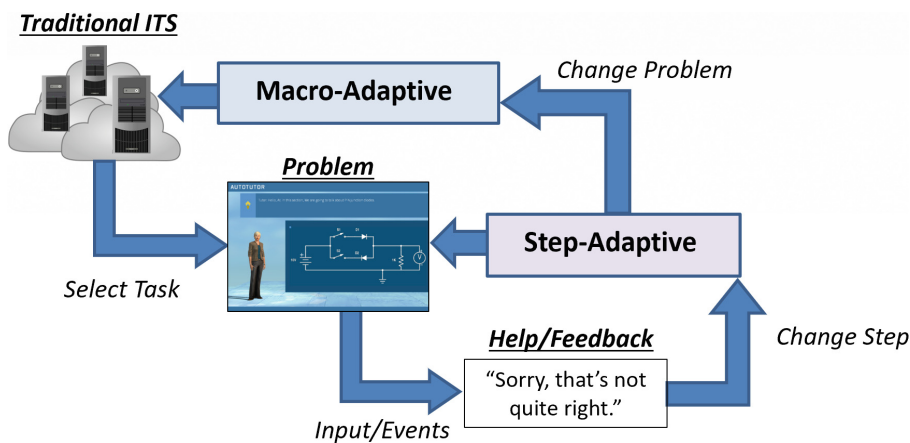


Fig. 1. Adaptivity at the step and problem levels that characterize AISs.

1.1 Successes of AISs

A substantial array of advanced AIS environments have sprung from the application of pedagogical and technological advancements to the field of CBT. Many of these have matured to the point of demonstrating significant learning gains over conventional instruction methods. These include efforts from concrete domains such as algebra and geometry, such as Cognitive Tutors [7–9] and ALEKS [10]. Systems targeting electronics (SHERLOCK [11], BEETLE-II [12]) and digital information technology (Digital Tutor [13]) have also proven effective.

Meta-analyses of intelligent tutoring systems (a subset of AIS leveraging artificial intelligence) have shown benefits compared to traditional classroom instruction or directed reading. The effects found varied substantially from a marginal $d = 0.05$ [14, 15] to a tremendous $d = 1.08$ [16], with most converging between $d = 0.40$ and $d = 0.80$ [17, 18]. This compares favorably to human tutoring, which can vary from roughly those same bookends depending on the expertise and effectiveness of the individual tutor [18]. Direct comparisons between human tutors and intelligent tutoring systems tend to show little or no statistical difference in learning outcomes [18, 19].

1.2 Conversational AIS

AISs that leverage verbal interaction can encourage more natural engagement with the subject matter [20, 21] and provide an avenue for addressing abstract domains. Explaining concepts extemporaneously urges learners to reflect and reason without relying on rigid repetition, formulas, or pre-defined answer options. Allowing learners to ask questions in return, through mixed initiative design, can focus conversational exchanges on areas or concepts with which the learner struggles or in which she has particular interest. Paralinguistic cues like pointing or facial expressions, instantiated with the use of embodied conversational agents, can increase realism and reinforce information by leveraging appropriate (and organically processed) gestures. Successful conversational AIS resources include AutoTutor [19, 21], GuruTutor [22], Betty's Brain [23], Coach Mike [24], Crystal Island [25], and Tactical Language and Culture System [26]. Each of these systems has demonstrated learning advantages relative to conventional instructional techniques. Conversational AISs have collectively addressed topics including computer literacy, physics, biology, and scientific reasoning.

This paper will focus on an application of AutoTutor. Some versions of AutoTutor include multiple conversational agents to provide more rich interaction dynamics. Conversational systems with multiple agents afford multiple interaction dynamics and greater interpersonal involvement [20, 27, 28]. In some iterations of AutoTutor, both a tutor agent and a peer agent engage in "trialogues" with the learner. This offers alternatives for presenting questions. The tutor agent can pose questions in a conventional instructor role, where the learner assumes the question-asker has the answer. Alternatively, the peer agent can pose the question from a position of relative weakness, asking the human learner for help. This tact can reduce pressure to provide full and complete information, or allow the human learner to play the role of a tutor. Also, novel formulations of answers may not be readily identifiable as correct or incorrect based on the pre-defined parameters constructed by experts. But having the peer agent rephrase the best match and submit as his or her own answer can avoid having the tutor agent erroneously evaluate learner responses as wrong. This avoids frustration and gives learners a chance to model more conventional expressions of the content.

1.3 Barriers to Adoption

Despite the learning gains demonstrated by the AISs above (both conversational and not), there remain several reasons that they have not gained ubiquity among the learning community. Primarily, they require considerable investment of resources and personnel. They require collaboration among many different specialties including experts in the domain of interest, computer science, and pedagogy. Additionally, the representation of information (e.g., graphical, interactive, conversational) likely requires specialists to synthesize expert input into a distinct approach. Smaller systems sometimes streamline this process by restricting the breadth of content available or depth of representation. In either case, creating broad swaths of content proves difficult.

This restriction of content naturally impedes widespread adoption. Any learning system entails some start-up costs where individuals learn how to interact with the system comfortably. Even exceptionally efficient systems would struggle to convince learners that the initial effort was justified for a mere subset of their learning goals. Evaluations then tends to focus on experimental groups to whom the AIS was assigned.

1.4 Hybrid Tutors

Although individual AIS content may suffer from lack of depth, breadth, or both, several independently produced systems may address the same domain, but represent slightly different (or complementary) content and engage in distinct pedagogical strategies. This array of technologies presents an opportunity to link existing learning resources centered around common subject matter. A confederated approach built into a common platform allows learners to have a common access point for a significantly enhanced range of content, presented in myriad ways. Such a system can include both adaptive and conventional learning resources, providing a full range of approaches to the material.

Through integration, AISs may overcome the single greatest hurdle to widespread adoption. This, in turn, could afford in-kind expansion of the individual learning resources based on the increased base of learners and subsequent opportunities for data-driven improvements. Comprehensive coverage of the target topic also provides invaluable learner data to human instructors at both the individual and classroom level. This makes a combined system ideal for classroom integration, where instructors can rapidly adjust lesson plans in response to automatically graded out-of-class efforts. A confederated system designed for use in conjunction with a human instructor is known as a hybrid tutor [29].

1.5 Meta Loop Adaptivity

Integrating systems together in a common platform creates new challenges to overcome. Learners may become overwhelmed by the amount of options, both in learning resources and in individual exercises or readings. Learners would naturally lack sufficient knowledge of the material and their understanding of it to navigate effectively, even without the addition of multiple representations. Human instructors understand progression through topics, but likely not the affordances and constraints of constituent learning resources. Further, this would not take full advantage of the capacity for learning technology to adapt to the individual. We must then construct a new “meta” loop of adaptivity to complement the existing within- and between-problem loops. Maintaining those two, a hybrid tutor must adapt to individual performance, history, and psychological characteristics at the level of learning resource. To do so requires two innovations: a way to translate progress among the several learning resources and a method of multifaceted assessment to begin recommendations.

Conversational interactions have unique advantages in providing the latter requirement.

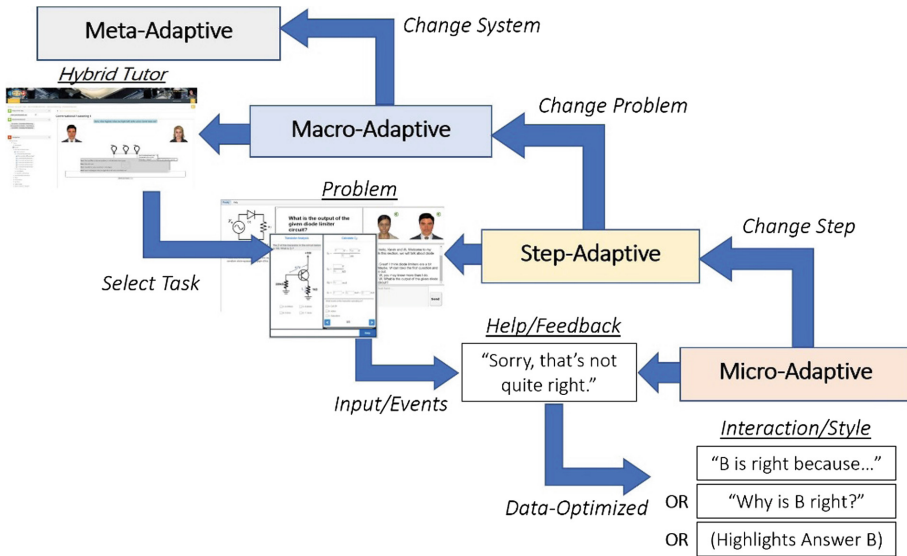


Fig. 2. Four-loop adaptivity, with AIS two-loop adaptivity plus meta-adaptive resource selection and micro-adaptive interaction style selection, utilized in AutoTutor.

2 ElectronixTutor

The potential benefits of confederated systems detailed above drove the creation of ElectronixTutor [29]. ElectronixTutor epitomizes the hybrid tutor approach, intended not to bypass classroom instruction, but to supplement it. The system addresses the topic of electricity and electrical engineering by leveraging diverse learning resources (both adaptive and static) in a unified platform. This platform (see Fig. 3) organizes content by

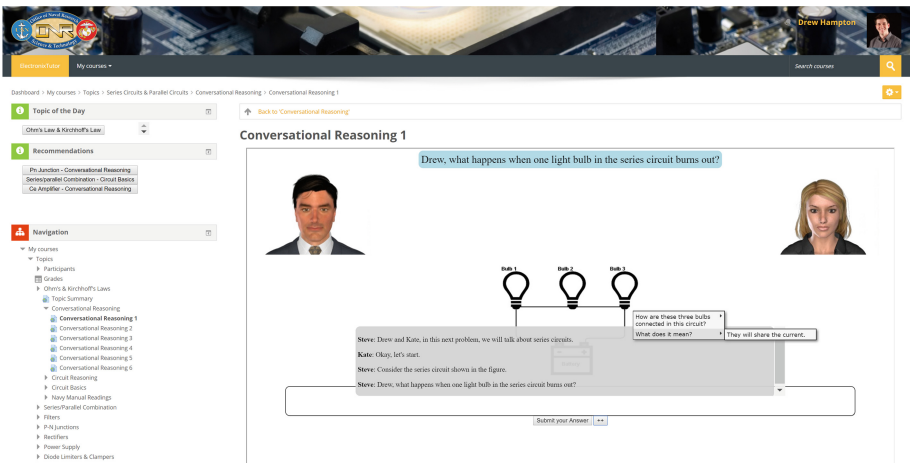


Fig. 3. The ElectronixTutor user interface, here showing an AutoTutor question with Point & Query engaged.

topic and resource, allowing learners to quickly find and engage with relevant problems or readings. Once selected, all resources appear in the activity window that occupies most of the screen, so the primary interface does not change for the learner.

These resources range from easy to difficult and employ a range of interaction types. As an introduction, *Topic Summaries* provide a brief (1–2 page) overview of each topic with hyperlinks to external resources like Wikipedia or YouTube. Built originally for the Navy, *ElectronixTutor* has a massive collection of static digital files from the *Navy Electricity and Electronics Training Series*. These comprehensive files are indexed by topic to ensure learners find the relevant sections quickly. Derived from a conversational AIS, *BEETLE-II* multiple-choice questions offer remedial exposure to important concepts represented in circuit diagrams. *Dragoon* circuit diagram questions, by far the most difficult resource, require learners to understand circuit components, relationships, variables, and parameters holistically. For mathematical rehearsal, *LearnForm* gives applied problems that break down into constituent steps, providing feedback along the way. Finally, *AutoTutor* provides conversational adaptivity for deep learning. I will discuss its functionality in depth below.

2.1 Knowledge Components

Critically, all individual learning resources contribute to a unified learner record store. This store translates progress among the many resources on several discrete levels. To accomplish this interfacing endeavor, *ElectronixTutor* conceptualizes learning as discrete knowledge components [30]. Knowledge components in *ElectronixTutor* (see Fig. 4) divide the target domain of electrical engineering into fifteen topics, device or circuit, and structure, behavior, function, or parameter, for a total of 120 knowledge components. Experts in electrical engineering use this structure to evaluate every problem and learning resource, indicating which knowledge component or components it addresses. That way, learner interactions with the problems produce a score from

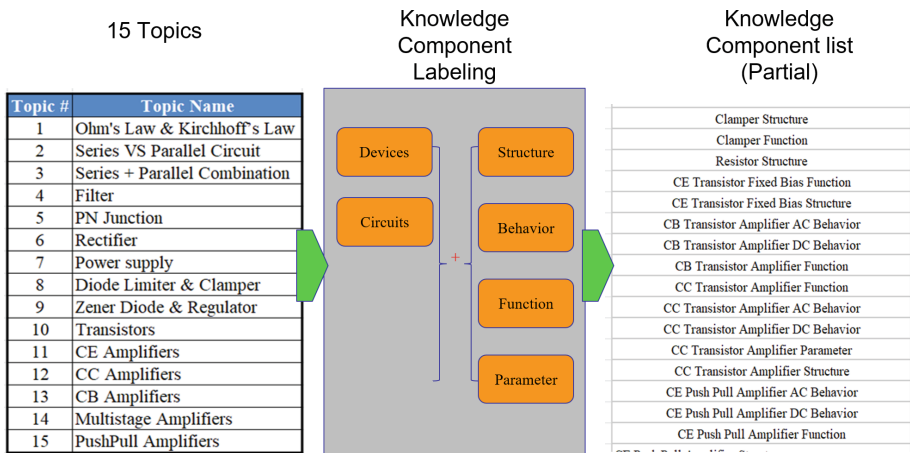


Fig. 4. Knowledge component mapping in *ElectronixTutor*

zero to one on each of the associated knowledge components that updates a comprehensive learner model appropriately.

2.2 Recommendations

Knowledge components and the associated learning record store enable the creation of an intelligent recommender system. The availability of copious amounts of information is bound to overwhelm if not presented in an order that makes sense to the individual learner. Recommendations must account for a learner's historical performance on a range of knowledge components, performance within learning resources, time since last interacting with content, and preferably psychological characteristics such as distinguishing between motivated and unmotivated learners.

ElectronixTutor offers two distinct methods of recommendation generation. One, instantiated in the *Topic of the Day* functionality, allows instructors to pick a topic either manually or through a calendar function. This restricts recommendations to content within that topic. The other, labeled *Recommendations* in Fig. 3, provides three options across the entire system. While the mechanisms for each differ slightly, both benefit from learners engaging with content that provides detailed feedback. Conversational AISs (AutoTutor in ElectronixTutor) surpass other available systems for several reasons.

2.3 AutoTutor

AutoTutor [19, 21] focuses on teaching conceptual understanding and encouraging deep learning. It accomplishes this by open conversational exchange in a mixed-initiative format. Availability of both a tutor agent and a peer agent opens the door to numerous conversational and pedagogical scenarios and techniques. These “trialogues” begin by introducing the topic and directing attention to a relevant graphical representation, calling the learner by name to encourage engagement.

Figure 5 demonstrates how AutoTutor (“Conversational Reasoning”) fits within the *Topic of the Day* recommendation decision tree. Here, the topic summary orients learners to the topic via brief, static introduction. This avoids asking questions on topics learners have not considered recently, thereby catching learners off their guard. From there, AutoTutor forms the hub of diagnosis and subsequent recommendation. It plays a similar role in populating the *Recommendations* section. It holds those position for several reasons.

Each full problem in AutoTutor has a main question with several components of a full correct answer. The conversational engine can extract partial, as well as incorrect, responses from natural language input. Assuming learners do not offer a full and correct answer upon first prompting, either the tutor agent or peer agent will follow up with hints, prompts, or pumps. This gives the learners every opportunity to express all of the information (or misconceptions) that they hold about the content at hand. In so doing, AutoTutor gathers far more than binary correct/incorrect responses. Learners who require multiple iterations to supply accurate information have a demonstrably more tenuous grasp of the concepts and their application. This provides invaluable diagnostic data regarding depth of understanding for the creation of intelligent recommendations.

Within AutoTutor, Point & Query [29] lowers the barrier for interacting with content by offering a simple mouse-over interaction with circuit diagrams. In conventional learning environments, students may fail to ask questions either because they do not know what appropriate questions are or they are intimidated by the scope of content presented. Point & Query removes that barrier by providing both the questions and the answers, orienting learners and providing baseline information. This greatly increases the absolute number of interactions with the learning program, and subsequently encourages more engagement by reinforcing question-asking behavior with immediate answers.

Further, main questions typically touch on several knowledge components simultaneously. AutoTutor’s focus on concepts naturally engenders questions about the relationships between multiple knowledge components. This both contributes to functional understanding and provides a breadth of information to the recommender systems. Presentation of multiple knowledge components in a single problem is not unique to AutoTutor. Notably, Dragoon requires extensive conceptual and technical knowledge of circuits to complete. However, finishing a problem in Dragoon is difficult

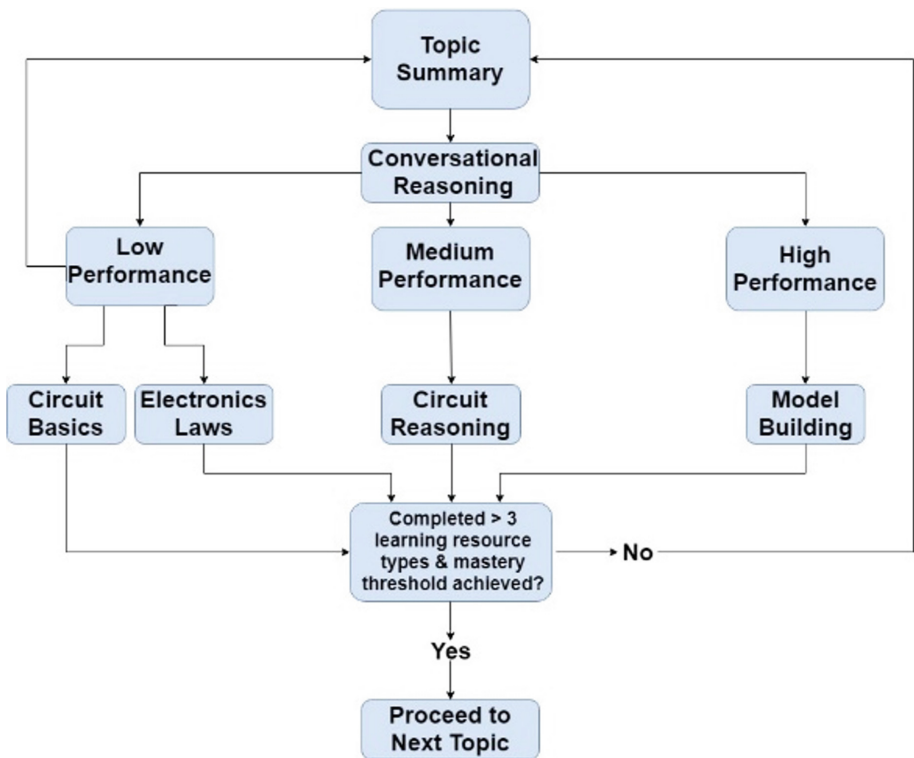


Fig. 5. A sample flow chart for ElectronixTutor intelligent recommendations within a topic designated by human instructors. Conversational reasoning (AutoTutor) plays a central role in diagnosis and subsequent recommendation.

without a significant amount of expertise in the content already. Dragoon struggles to walk learners through problems unless they are already at a high level. AutoTutor, by contrast, provides ample lifelines for struggling learners. Point & Query establishes a baseline of information needed to answer the question, then hints, prompts, and pumps provide staged interventions, eliciting knowledge and encouraging reasoning to give every opportunity for learners to complete the problem.

Open-ended interactions also have the benefit of providing opportunities to assess more psychological characteristics than rigidly defined learning resources. Willingness to continue engaging with tutor and peer agents through more than a few conversational turns indicates that a learner is motivated to master the content. These learners may prefer to push the envelope rather than refreshing past material. Long answers indicate verbal fluency, perhaps a suggestion that they will respond well to more conversational instruction moving forward. Evaluation on these lines can determine not just what problems should come next (the macro-adaptive step), but also how to respond to learner input within a problem (micro-adaptivity—see Fig. 2).

3 Conclusions

AISs represent a substantial advancement over conventional computer-based training. Providing adaptivity within or between problems allows learners to have more personalized learning experiences. However, the trade-off between depth of instruction and investment to create it means that most individual systems lack the breadth to engender widespread adoption. Integration of these systems into a common platform serves two important purposes. First, it constitutes a method of improving individual learning in AISs by making multiple representations of content available and broadening the total area covered. Second, it represents a practical method of encouraging more learners and educational institutions to use it, providing data for improvement and incentive to continue creating.

However, fundamental challenges must be overcome to realize this idea. First, some method of translating progress and collecting learner characteristics must be implemented. Second, the system must include intelligent recommendations for how to proceed, thus avoiding overwhelming learners with too many unorganized options. ElectronixTutor approaches the first problem through knowledge components that discretize the domain relative to individual learning resources. The second problem requires a learning resource with exceptional diagnostic power to offer a preliminary assessment.

Conversational reasoning questions serve this function, in this case through the AutoTutor AIS. These questions probe depth of understanding on a conceptual level in a way that few AISs can. Further, the inclusion of multiple knowledge components in a single interaction allows for considerable breadth. An array of remedial affordances allows fine-grained evaluation of knowledge and encourage completion despite potential difficulties. These embedded resources include Point & Query functionality to encourage initial interaction with content, along with a series of hints, prompts, and pumps to get learners across the finish line. Finally, the open-ended interaction provides an opportunity to glean psychological characteristics that can inform recommendations

beyond that made possible with performance history alone. For these reasons—depth of assessment, breadth of content covered, availability of remedial steps, and potential for psychological assessment—conversational AISs form the cornerstone of intelligent recommendations in hybrid tutors.

References

1. Sottolare, R., Brawner, K.: Exploring standardization opportunities by examining interaction between common adaptive instructional system components. In: Proceedings of the First Adaptive Instructional Systems (AIS) Standards Workshop, Orlando, Florida (2018)
2. Graesser, A.C., Hu, X., Nye, B., Sottolare, R.: Intelligent tutoring systems, serious games, and the Generalized Intelligent Framework for Tutoring (GIFT). In: Using Games and Simulation for Teaching and Assessment, pp. 58–79 (2016)
3. Wainer, H., Dorans, N.J., Flaughner, R., Green, B.F., Mislevy, R.J.: Computerized Adaptive Testing: A Primer. Routledge, New York (2000)
4. VanLehn, K.: The Behavior of tutoring systems. *Int. J. Artif. Intell. Educ.* **16**, 227–265 (2006)
5. Sottolare, R., Graesser, A.C., Hu, X., Holden, H. (eds.): Design Recommendations for Intelligent Tutoring Systems. Learner Modeling, vol. 1. U.S. Army Research Laboratory, Orlando (2013)
6. Woolf, B.P.: Building Intelligent Interactive Tutors. Morgan Kaufmann Publishers, Burlington (2009)
7. Alevan, V., McLaren, B.M., Sewall, J., Koedinger, K.R.: A New paradigm for intelligent tutoring systems: example-tracing tutors. *Int. J. Artif. Intell. Educ.* **19**(2), 105–154 (2009)
8. Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.: Intelligent tutoring goes to school in the big city. *Int. J. Artif. Intell. Educ.* **8**, 30–43 (1997)
9. Ritter, S., Anderson, J.R., Koedinger, K.R., Corbett, A.: Cognitive tutor: applied research in mathematics education. *Psychon. Bull. Rev.* **14**, 249–255 (2007)
10. Falmagne, J., Albert, D., Doble, C., Eppstein, D., Hu, X.: Knowledge Spaces: Applications in Education. Springer, Heidelberg (2013). <https://doi.org/10.1007/978-3-642-35329-1>
11. Lesgold, A., Lajoie, S.P., Bunzo, M., Eggan, G.: SHERLOCK: a coached practice environment for an electronics trouble-shooting job. In: Larkin, J.H., Chabay, R.W. (eds.) Computer Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches, pp. 201–238. Erlbaum, Hillsdale (1992)
12. Dzikovska, M., Steinhauer, N., Farrow, E., Moore, J., Campbell, G.: BEETLE II: deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *Int. J. Artif. Intell. Educ.* **24**, 284–332 (2014)
13. Fletcher, J.D., Morrison, J.E.: DARPA Digital Tutor: Assessment data (IDA Document D-4686). Institute for Defense Analyses, Alexandria, VA (2012)
14. Dynarsky, M., et al.: Effectiveness of reading and mathematics software products: Findings from the first student cohort. U.S. Department of Education, Institute of Education Sciences, Washington, DC (2007)
15. Steenbergen-Hu, S., Cooper, H.: A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *J. Educ. Psychol.* **106**, 331–347 (2013)
16. Dodds, P.V.W., Fletcher, J.D.: Opportunities for new “smart” learning environments enabled by next generation web capabilities. *J. Educ. Multimed. Hypermedia* **13**, 391–404 (2004)
17. Kulik, J.A., Fletcher, J.D.: Effectiveness of intelligent tutoring systems: a meta-analytic review. *Rev. Educ. Res.* **85**, 171–204 (2015)

18. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. *Educ. Psychol.* **46**, 197–221 (2011)
19. Graesser, A.C.: Conversations with AutoTutor help students learn. *Int. J. Artif. Intell. Educ.* **26**, 124–132 (2016)
20. Johnson, W.L., Lester, J.C.: Face-to-face interaction with pedagogical agents, Twenty years later. *Int. J. Artif. Intell. Educ.* **26**(1), 25–36 (2016)
21. Nye, B.D., Graesser, A.C., Hu, X.: AutoTutor and family: a review of 17 years of natural language tutoring. *Int. J. Artif. Intell. Educ.* **24**(4), 427–469 (2014)
22. Olney, A.M., Person, N.K., Graesser, A.C.: Guru: designing a conversational expert intelligent tutoring system. In: *Cross-Disciplinary Advances in Applied Natural Language Processing: Issues and Approaches*, pp. 156–171. IGI Global (2012)
23. Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., Roscoe, R.: Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Res. Pract. Technol. Enhanc. Learn.* **5**, 123–152 (2010)
24. Lane, H.C., Noren, D., Auerbach, D., Birch, M., Swartout, W.: Intelligent tutoring goes to the museum in the big City: a pedagogical agent for informal science education. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) *AIED 2011. LNCS (LNAI)*, vol. 6738, pp. 155–162. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21869-9_22
25. Rowe, J.P., Shores, L.R., Mott, B.W., Lester, J.C.: Integrating learning, problem solving, and engagement in narrative-centered learning environments. *Int. J. Artif. Intell. Educ.* **21**, 115–133 (2011)
26. Johnson, L.W., Valente, A.: Tactical language and culture training systems: using artificial intelligence to teach Foreign languages and cultures. *AI Mag.* **30**, 72–83 (2009)
27. Craig, S.D., Twyford, J., Irigoyen, N., Zipp, S.A.: A Test of spatial contiguity for virtual human’s gestures in multimedia learning environments. *J. Educ. Comput. Res.* **53**(1), 3–14 (2015)
28. Graesser, A.C., Li, H., Forsyth, C.: Learning by communicating in natural language with conversational agents. *Curr. Dir. Psychol. Sci.* **23**, 374–380 (2014)
29. Graesser, A.C., et al.: ElectronixTutor: an adaptive learning platform with multiple resources. In: *Proceedings of the Interservice/Industry Training, Simulation, & Education Conference (IITSEC 2018)*, Orlando, FL (2018)
30. Tackett, A.C., et al.: Knowledge components as a unifying standard of intelligent tutoring systems. In: *Exploring Opportunities to Standardize Adaptive Instructional Systems (AISs) Workshop of the 19th International Conference of the Artificial Intelligence in Education (AIED) Conference*, London, UK (2018)