An incremental approach for building an intelligent tutoring system using a dialog based agent

Siddharth Kanungo

Table of contents

Abstract A text-based conversational agent would allow guided learning to be more personalized and adaptive to the needs of learner. Part of the solution involves creating dialog scripts to be used by agents to instruct the learner about a subject and part of the solution involves developing an interface through which the agent interacts with the learner. Such a system enables the learner to enrich their personal knowledge database and aids in retaining their learning through spaced repetitions. Contrary to popular beliefs, the conversational agent can be an effective tool for learning even without possessing the ability to process natural language. We have built a tutoring system using a dialog based agent called Primer and we state the rationale and architecture of such a system in this paper.

1 Introduction

The traditional methods for guided learning, such as textbooks, recorded lectures, and instructor-led learning, are insufficient in the dissemination of the vast amount of scientific knowledge that is continuously being generated every year. The idea of building a conversational agent stems from the need that access to the collective scientific knowledge should be made easier. Reasons for the inability to access such large amount to knowledge may be attributed to pre-requisites debt, inability to scale traditional methods of learning, inadvertent denial of access, high opportunity costs associated with the pursuit of acquiring knowledge, a high perceived barrier to entry and the issue of self-discipline coupled with low attention span of learners.

Guided learning in the form of a conversational agent will be able to mitigate such issues. Apart from that, we look into different problems for which guided learning can be an effective

tool.

1.1 Bloom's 2 Sigma Problem

In 1984, educational researcher Benjamin Bloom published his findings[2] that using mastery learning, one-on-one tutoring led to a two-sigma improvement in student's performance over the students who were instructed using conventional methods. In the paper, he posed the question:

> Can researchers and teachers devise teaching-learning conditions that will enable the majority of students under group instruction to attain levels of achievement that (at present) can be reached only under good tutoring conditions?

This question is referred as the **Bloom's 2 Sigma problem**. Bloom saw his key factor influencing his findings: one-on-one tutoring is too costly for most societies to bear on a large scale.

Also a study regarding the effectiveness of direct instruction [12] shows that the effects of direct instruction using mastery learning are large. The paper notes that

> Yet, despite the very large body of research supporting its effectiveness, DI has not been widely embraced or implemented

Due to technological advances, the idea of one-on-one tutoring without a human actor is feasible with the aid of a conversational agent. Although, it's effectiveness compared to human actors led tutoring needs to be improved continuously.

1.2 Gatto's Critisicm of Schooling

John Taylor Gatto in his book Dumbing us down [4] notes that the current widespread schooling system inadvertently teaches the following:

1. **Confusion**: Teaching out of context, fragmentation, facts loosely or not at all connected to life or other pieces of information.

2. **Indifference**: It encourages superficial and shallow understanding of a wide spectrum of subjects rather than deep knowledge of one. The classes are structured based on time allotments and the students are not provided enough time to delve deeply into a subject. Gatto notes that when it's time to move to another class, the student must stop caring or learning about xyz subject and now to focus on abc.

3. **Class position**: Students are restricted in their assigned class.

4. **Emotional & intellectual dependency, provisional self-esteem**: Stars, grades, prizes, and honors are often awarded based on the ability of the student for doing what they are told. The expert chooses what to study and who has done a good job.

5. **Surveillance**: Schools afford no privacy for students. The movements of students in schools are highly restricted.

Gatto states that the goal of the school is compliance, not curiosity. Similar concerns have been raised by Ivan Illich [7] and Pawan Gupta[5]. Ivan Illich suggests creation of Learning Webs to connect learners with "masters" and Pawan Gupta advocates that the schooling should be relatable to the local environments.

In order to counter against such schooling, Gatto advocates homeschooling where the learners are in charge of their own learning. Although homeschooling has never been easier at this point in time in history due to technological advances, its widespread adoption is limited.

Guided-learning in the form of a conversational agent will be an effective tool for parents to home school their children. The conversational agent can be extended to include local context into the learning lesson so that they students can map their understanding about the world from their local environments.

1.3 Deep Work

In his book, Deep Work [9], computer science professor, Dr.Cal Newport defines deep work as

> Professional activities performed in a state of distraction-free concentration that pushes your cognitive capabilities to their limit. These efforts create new value, improve your skill, and are hard to replicate.

An undertaking of learning a subject by oneself requires the ability to do deep work frequently over a long period of time. Barbara Oakley in her book Mindshift[10] notes that

> Becoming an expert in something new, whatever the subject, means building small chunks of knowledge using day-by-day practice and repetitions. Gradually, these small chunks can be knit together into mastery.

Although deemed revolutionary, Massive Online Organized Courses (MOOCs) have failed to entice most of the students to complete the courses which they had signed up for. This signifies that the obstacle towards self-learning subjects is more affiliated with the self-discipline than the availability of learning materials. This also posits that the responsibility of the tutor is not only to equip the students with the information about the subject but also ensuring mitigation of the known issues pertaining to self-discipline among the learner.

A conversational agent will help in eliminating a lot of friction in the undertaking by providing digestible chunks of new information, helping in retaining information learned and acting as a personal knowledge database & assistant.

1.4 A Mathematician's Lament

In his article A Mathematician's Lament[8], mathematician Paul Lockhart notes the following:

> If I had to design a mechanism for the express purpose of destroying a child's natural curiosity and love of pattern-making, I couldn't possibly do as good a job as is currently being done. I simply wouldn't have the imagination to come up with the kind of senseless, soul-crushing ideas that constitute contemporary mathematics education.

The pedagogy of mathematics warrants its careful treatment as an art. Mathematical subjects require that the student approach them as a creative pursuit rather than being dependent on rote learning of formulas and pattern matching. This requires that mathematical topics are needed to be experimented with and improved continuously by gauging their effectiveness at elevating the learner's mathematical aptitude.

The issue of incorrect pedagogy is also prevalent in other fields of studies as well. And experimenting with the curriculum and the way of introducing new information will result in improvements.

Traditional methods of guided learning have less scope for improvement and less space for experimentation. However, the script and features of the conversational agent can be continuously improved based on feedback resulting in more effectiveness.

1.5 The Eliza Effect

In 1966, Computer Scientist Joseph Weizenbaum, created a chatbot program called ELIZA. When executing Weizenbaum's DOCTOR script, ELIZA parodied a psychotherapist, largely by rephrasing the *patient's* replies as question. ELIZA's DOCTOR script was found to be surprisingly successful in eliciting emotional responses from users who, in the course of interacting with the program, began to ascribe understanding and motivation to the program's output. Upon observation, researchers discovered users unconsciously assuming ELIZA's questions implied interest and emotional involvement in the topics discussed, even when they consciously knew that ELIZA did not simulate emotion.

In his book *Fluid concepts and creative analogies*[6], Douglas Hofstadter refers to this phenomenon as the ELIZA Effect. He defines it as

> The susceptibility of people to read far more understanding than is warranted into strings of symbolsespecially wordsstrung together by computers.

A trivial example of the specific form of the Eliza effect, given by Douglas Hofstadter, involves an automated teller machine which displays the words "THANK YOU" at the end of a transaction. A casual observer might think that the machine is actually expressing gratitude; however, the machine is only printing a preprogrammed string of symbols.

We developed a conversational agent called Primer and in this paper we outline the conceptual framework with the underlying rationale, principles and architecture of the system. This paper challenges the notion that an effective conversational agent will require proficient ability to process natural language. We propose that an effective platform with the agent having no ability to understand natural language can be built by taking advantage of the ELIZA Effect and incrementally improving its ability to understand basic user intent. The goal of building such a system is to provide all those who want to learn with access to an tutor for teaching them at any time in their lives.

2 Conceptual Framework

2.1 Objectives

We define the basic objectives of our intelligent tutoring system as follows:

- the system should be able to instruct a learner regarding a subject in a dialog based format

- the system should be able to provide enough learning aids to the learner to deepen the understanding of the subject

- the system should be able to assist the learner in revising subject for maximum retention

In order to understand the conceptual framework for building such a tutoring system, we will begin by basing our system on a simplified model of a classroom instructor. We will assume that the way this particular instructor instructs is by following a pre-defined format. We also assume that this particular instructor does not possess any art of storytelling to captivate the students. This particular assumption helps us to base our model such that the lectures consists only of pure information directly associated with the lectures and not any trivial buffer to keep students engaged. We call this system, the unimaginative instructor. We will also assume that the system is teaching is only one student.

2.2 The blueprint of an unimaginative instructor

In order to teach a student a particular course, the unimaginative instructor prepares a curriculum or customizes a widely used one. A curriculum dictates the relative ordering of the course topics for better understanding. The instructor divides the course into sections according to the curriculum and distributes them over a number of lectures.

The standard template used by the Unimaginative Instructor can be as follows:

1. States the topic

2. Provides historical background and importance of the topic

3. Shows applications of the topics of the topic

4. Explain basic concepts and how it fits together with the rest of the course topics

5. Shows the derivation of equations and formulaes, if any.

6. Demonstrates the usage of such equations

7. Moves to step 1 for the next topic

The Unimaginative Instructor does the job of conveying of ideas necessary for learning the course pretty well. However, the course comprehension of the student for the course may be low as the student may find the lecture dull and hard to pay attention to.

We can improve students comprehension by introducing interactivity to the instructor.

2.3 Conversational Tutor

In order to add interactivity, we will convert the lectures from the unimaginative instructor into text format. The lectures would be classified into the following:

- **Information**: A formal presentation of main ideas pertaining to the topic

- **Interaction**: An interaction is a question or prompt posed to the learner by the system which requires learner input. An interaction consists of:

  1. **Conversation**: A short informal and often personalized discussion regarding the topic. In Conversation, the learner could interact with the user using predefined replies or an input text area.

  2. **Widget**: An interactive teaching aid. Widgets consist of multiple choice questions, reordering, fill in the blanks, match the following, an interactive chessboard, an explorable etc.

```
Primer : Nell, what is the answer to the ultimate question of life, the
         universe, and everything?
Primer : Options are [a]Happiness  [b]42  [c]108  [d]666
Nell   : It is [b]42, of course.
Primer : Wrong. It is [a]Happiness and don't let any sci-fi nut tell you
         otherwise.
```

Listing 1: Learning Atom with Interaction using MCQ Widget with the learner Nell

Interaction helps in retaining the student's attention and forces the student to attempt to utilize information acquired from the course to submit inputs to the questions posed during in the interaction.

Each topic is broken down in sets of Interaction and/or Information. Such a set of Interaction and/or Information is named **Learning Atom** for convenience.

The main ideas of the topic are communicated using Information, while the Interaction with the learner is done with the help of Conversation and Widgets.

Listing 1 shows a Learning Atom with Interaction using MCQ Widget. Listing 2 shows a Learning Atom with Interaction using *Conversation*.

The goal of each Learning Atom is to increase and test learner's understanding of the subject. In order to do that each Learning Atom is associated with a learning objective based on Revised Blooms Taxonomy.

2.4 Revised Blooms Taxonomy

In 1956, Benjamin Bloom headed a group of educational psychologists who developed a classification system for levels of cognitive skills and learning behaviour. The classification system they created is often referred to as Bloom's Taxonomy[3].

In 2001, Lorin Anderson and a group of cognitive psychologists, curriculum theorists and instructional researchers, and testing and assessment specialists published a revision of Blooms Taxonomy with the title A Taxonomy for Teaching, Learning, and Assessment[1].

A statement of a **learning objective** contains a **verb** (an action) and an **object** (usually a noun).

- The **verb** generally refers to [actions associated with] the intended **cognitive process**.
- The **object** generally describes the **knowledge** students are expected to acquire or construct. (Anderson and Krathwohl, 2001, pp. 4–5)

In this model, each of the colored blocks shows an example of a learning objective that generally corresponds with each of the various combinations of the cognitive process and knowledge dimensions.

Remember: these are **learning** *objectives*—not learning *activities*. It may be useful to think of preceding each objective with something like: "Students will be able to . . ."

*Anderson, L.W. (Ed.), Krathwohl, D.R. (Ed.), Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives (Complete edition). New York: Longman.

Model created by: Rex Heer
Iowa State University
Center for Excellence in Learning and Teaching
Updated January, 2012
Licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
For additional resources, see:
www.celt.iastate.edu/teaching/RevisedBlooms1.html

IOWA STATE UNIVERSITY
Center for Excellence in
Learning and Teaching

Figure 1: A 3d representation of Revised Bloom's Taxonomy. Credits to Iowa State University

The Revised Bloom's Taxonomy provides a valuable framework to develop performance tasks, creating questions, constructing problems and helps in assessment of the learner's ability. The Taxonomy classifies the learning behaviour into two dimensions:

1. Cognitive Process Dimension

2. Knowledge Dimension

Cognitive Process Dimension The cognitive process dimension represents a continuum of increasing cognitive complexity from remember to create. Anderson and his group identified 19 specific cognitive processes that further clarify the bounds of the six categories.

1. **Remember**: recognizing (*identifying*), recalling (*retrieving*)

2. **Understand**: interpreting (*clarifying, paraphrasing, representing, translating*), exemplifying (*illustrating, instantiating*), classifying (*categorizing, subsuming*), summarizing (*abstracting, generalizing*), inferring (*concluding, extrapolating, interpolating, predicting*), comparing (*contrasting, mapping, matching*), explaining (*constructing models*)

3. **Apply**: executing *(carrying out)*, implementing *(using)*

4. **Analyze**: differentiating *(discriminating, distinguishing, focusing, selecting)*, organizing *(finding, coherence, integrating, outlining, parsing, structuring)*, attributing *(deconstructing)*

5. **Evaluate**: checking *(coordinating, detecting, monitoring, testing)*, critiquing *(judging)*

6. **Create**: generating *(hypothesizing)*, planning *(designing)*, producing *(construct)*

Knowledge Dimension The knowledge dimension represents a range from concrete (factual) to abstract (metacognitive).

1. **Factual**

   - knowledge of terminology
   - knowledge of specific details and elements

2. **Conceptual**

   - knowledge of classifications and categories
   - knowledge of principles and generalizations
   - knowledge of theories, models, and structures

3. **Procedural**

   - knowledge of subject-specific skills and algorithms
   - knowledge of subject-specific techniques and methods
   - knowledge of criteria for determining when to use appropriate procedures

4. **Metacognitive**

   - strategic knowledge
   - knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
   - self-knowledge

A integrated representation of cognitive and knowledge dimension and how the action verb can be used to frame learning objective is demonstrated in Fig 1

2.5 Learning Atom Objective

In order to frame the learning objective for the Learning Atom, a suitable verb (intended action) is matched with a suitable objective with the help of the Revised Bloom's Taxonomy.
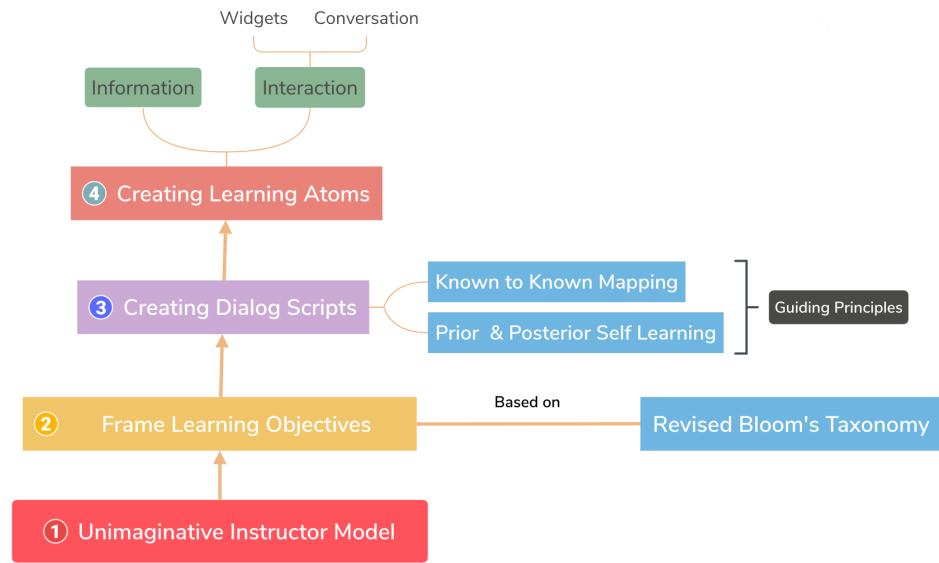
Figure 2: The conceptual framework for building course scripts for Primer

```
Primer : Nell, Do you recall the mnemonic for naming the carpal bones?
Nell   : Yes, I do.
Primer : Can you list the mnemonic?
Nell   : Some Lovers Try Positions That They Can't Handle
Primer : The mnemonic for remembering Carpal bones is "Some Lovers Try
         Positions That They Can't Handle"
Primer : Can you map the mnemonic to the bones?
Nell   : Some−Schapoid, Lovers−Lunate, Try−Triquetrum, Positions−Pisiform,
         That−Trapezium, They−Trapezoid, Can't−Capitate and Handle−Hamate
Primer : Excellent.
```

Listing 2: Learning Atom with learning objective - To List the carpal bones

Listing 1 demonstrates a learning atom with the learning objective of naming the carpal bones. In order to formulate such integration into Interaction so that the learning objective is achieved, there are two guiding principles.

- Prior and Posterior Self Learning

- Known to Known Mapping

2.6 Principle of Prior and Posterior Self Learning

The central idea of conversational learning is to implement Prior and Posterior Learning (PPL). In Prior and Posterior Learning, the agent questions the learner about his prior understanding about a certain topic. After the learner replies his own ideas about the topic, the agent

```
Primer  :  Nell, what do you think is a hypothesis?
Nell    :  I think a hypothesis is a proposed explanation for a phenomenon.
Nell    :  For a hypothesis to be a scientific hypothesis, the scientific
           method requires that one can test it.
Primer  :  Simply, hypothesis is an educated guess.
Primer  :  For a hypothesis to be scientific, it has to be a specific,
           testable prediction.
Primer  :  It describes in concrete terms what you expect will happen in a
           certain circumstance.
```

Listing 3: Learning Atom with the user having correct assumption

```
Primer  :  Nell, what do you think is a hypothesis?
Nell    :  I think it is an educated guess.
Primer  :  Simply, hypothesis is an educated guess.
Primer  :  For a hypothesis to be scientific, it has to be a specific,
           testable prediction.
Primer  :  It describes in concrete terms what you expect will happen in a
           certain circumstance.
```

Listing 4: Learning Atom with the user having incorrect assumption

replies with the accepted answer about the topic. The learner can self-evaluate his/her own understanding by comparing against the accepted answer provided by the agent. This lets the user to understand the gaps present in his understanding and accepted understanding in case he is way off. In case the user's understanding of the topic is correct, this reinforces confidence in the learner.

It can be noted that the agent doesn't offer *correct* answer rather an accepted answer. It is to reinforce the idea that the agent (authority) can be wrong and so can be the accepted answer.

An example of such Learning Atom is shown is Listing 4 with the learner Nell who is incorrect in her assumption and an example of such learning where the learner Nell is correct is shown in Listing 3.

2.7 Principle of Known to Known Mapping

The other principle used to achieve the learning objective of the Learning Atom is named Principle of Known to Known Mapping. It is based on the idea that every topic is built on the foundations of ideas resulting from one or more sets of information. The role of the instructor is often to demonstrate how the topic can be derived from its reference information. Demonstrating such instructing without keeping the learner engaged may result in the learner passively absorbing ideas that may or may not map into the learner's knowledge database.

In Known-to-Known Mapping, the instructor provides all necessary information as ingredients to derive the required information to the learner and the learner uses the available

```
Primer  :  Nell, how does the water flow?
Nell    :  Primer, water flows from higher altitude to lower altitude.
Primer  :  Correct Nell. Now, how does the river Ganges flow in India?
Nell    :  River Ganga, starts from Himalayas passess through the states
           like Uttar Pradesh, Himachal Pradesh, Bihar, West Bengal and then
           finally merges with the Bay of Bengal.
Primer  :  If that's the case, what can you comment about the topographical
           nature of India.
Nell    :  The northern parts of India should be at higher elevation and
           the slope must gradually lower in the middle part towards the
           eastern side of the country.
```

Listing 5: Learning Atom using the principle of Known to Known Mapping

```
Primer  :  Nell, can you describe the topographical nature of India.
Nell    :  Nope, Primer. I don't think so.
Primer  :  Well Nell. The northern part of India are at higher altitude and
           the altitude tends to decrease as we move towards the southern
           part of the country.
Nell    :  Okay Primer. :|
```

Listing 6: Learning Atom using the principle of Unknown to Known Mapping

ingredients to infer the topic's central idea. This is an example of active learning by the learner. This principle is hard to replicate in a live instructor-led classroom due to the constraint of time and in a recorded lecture due to the lack of interactive input.

A conversational learning agent implements the principle of known-to-known mapping in the following way:

1. The agent poses questions about the related concept that it believes the learner is aware of

2. The agent asks the learner to relate the concepts and hints towards the topic objective

3. The learner may or may not be able to derive the topic but a sincere effort results in deeper understanding.

Such a system ensures that the new knowledge is added to the system maps to something that the learner is aware of and enriches the knowledge database. An example of the principle in action is shown in the Listing 5. Here, the learner derives the required information from previously acquired knowledge. The same learning objective is demonstrated using the principle of Unknown-to-Known Mapping in the Listing 6. Here, the learner acquires a new piece of knowledge without linking it with exisiting learner's database.

A dialog script is generated based upon the above two principles to achieve the learning objectives of each Learning Atoms. The script is used by the conversational agent to instruct the learner.
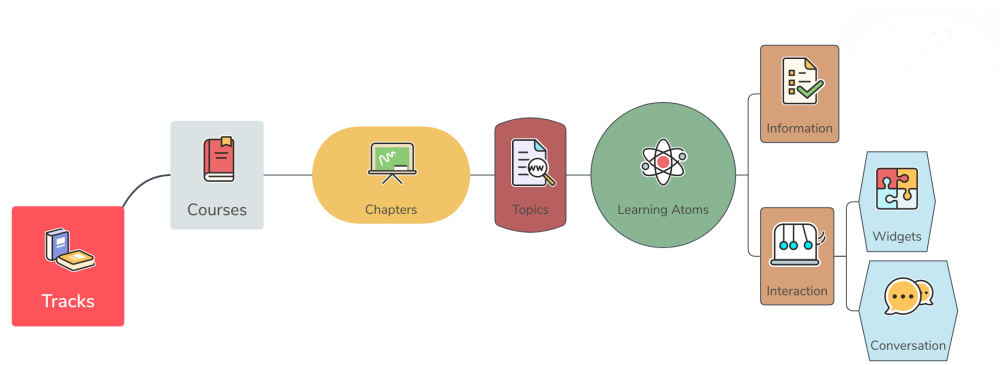
Figure 3: Framework for creation of dialog scripts for Primer

2.8 Framework of Primer Courses

For creating courses as dialog scripts for Primer, the subject of learning is divided into *Tracks*. An example of Track is *Machine Learning*. Tracks are further divided into courses. For example Machine Learning Track can divided into courses as *Regression, Support Vector Machines etc*. Each course is further divided into *Chapters* and *Chapters* are divided into *Topics*.

Each Topic is divided into sets of Learning Atoms with a specific learning objective based on Revised Bloom's Taxonomy. Learning Atoms are consist of Information and Interaction. This framework is shown in the figure 2 and 3.

This is a basic conceptual framework of the Intelligent Tutoring System, *Primer*.

Although the tutoring system is ready to instruct the learner, it needs to be extended for a better understanding of the learner.

3 Extending the Tutoring System

3.1 Summarizing Prompt

At the end of each topic, the learner is provided with an opportunity to explain their understanding about the recently learned topic. This helps the learners in formulating their understanding of concrete statements in their own language. This learner created summary is also added to an automatically generated notebook at the end of the course.

If a learner is facing difficulty into explaining the concepts and information, then it is taken into account by the feedback system. How well the dialog script for the topic is able to achieve its learning objective is tracked with Explainability Metric.

3.2 Explainability and Dialog Script as a Software

The *Explainability* of the course refers to the fraction of students who took the course and were not able to understand the topic to the total number of learners who completed the course.

This is provided by the students in the feedback system which also asks the students to pinpoint which topic was more difficult to understand too. Feedbacks without reasoning is

$$E = 1 - \frac{n_s}{n_t} \tag{1}$$

where $E =$ Explainability of the topic,

$n_s =$ number of learners find it lacking to explain the intended topic

$n_t =$ total number of learners providing feedback

Figure 4: Calculation of Explainability

not taken into consideration for calculating explainability.

Low explainability for a topic demands the course script to modified and improved. This posits the course script to be akin to software which improves itself on every release and is versioned accordingly. The topic which has low explainability can be modified to be divided into smaller parts to facilitate better understanding and the usage of creativity to provide understanding.

In order to deepen the understanding, the course often carries a project component, which the learner needs to complete before the learner can complete the course.

3.3 Project

The project component of the tutoring system is a tools agnostic dialog script that underlies a real-world problem statement and uses the newly acquired course concepts to solve them. With the help of the project, the agent guides the learner in understanding the application of the topic in the problems. Example of such a project is building a Convolutional Neural Network from scratch and using it for image classification for a real-world dataset.

3.4 Learner Notes

At any point during the topic or during revision, the learner has the ability to add notes and questions.

The there system allows different forms of notes:

1. **Topic Notes**: Notes written during the course

2. **Revision Notes**: Notes written at the time of revision

3. **Questions**: Questions that arise in the mind of the learner

4. **Equation Notes**: Notes written against an equation

Additionally, all the terms and definition introduced in the topic are compiled together for easier access of the learner. The learner can also coin new terms and its definition for better retention.

3.5 Learner Questions

The Tutoring system also includes a community forum where the learner can post their questions. After a certain interval, the questions are compiled and added to the dialog scripts in their next improved version. The tutoring system also helps the learner to document their answer for their questions.

3.6 Automatically Generated Notebooks

At the successful completion of course, the learner can generate a personalized course notebook in the portable document format(PDF). This document contains lecture notes, learner's notes, learner's responses, terms, and questions.

This is provided to free the subject of any compulsion during the course to scribble down on paper and provide utmost attention to the topic being taught. The notebook also acts a course journal which can be revisited even after a period of time and will be easier to understand due to its personalized nature.

3.7 Revision and Auto Generation of Flash Cards

User Notes

```
The actual value of the covariance is not meaningful. It's affected by
the scale of the two variables. That is why we calculate the
correlation coefficient. In order to make something interpretable from
the covariance information.
```

Front Side of the Flash Card

```
The actual value of the _____ is not meaningful. It's _____ by the
_____ of the two variables. That is why we calculate the correlation
_____. In order to make something _____ from the covariance
information.
```

Back Side of the Flash Card

```
The actual value of the covariance is not meaningful. It's affected by
the scale of the two variables. That is why we calculate the
correlation coefficient. In order to make something interpretable from
the covariance information.
```

Listing 7: Automatic Generation of Flash Cards from User Notes

The responsibility of the tutoring system also includes that the learner is able to retain the topics learned during course. After completion of any topic, the learner can revise the topic by going through the conversation with the agent regarding the topic. Also after the topic is completed, predefined flashcards relevant to the topic are added to the revision deck. Apart from that, flash cards are generated from the learner's notes, completed questions and terms. The flashcards generated from user notes is done by first dividing the text into a string of separate sentences and then eliminating the stop words from the sentence. Stop words are the words in the language which appear frequently but doesn't provide additional meaning to the sentence. An example of the autogenerated flash card from the user notes can be seen in

the Listing 7.

3.8 Retention

The responsibility of the agent includes ensuring that the subject retains the course in the long term. In order to do that, there are two kinds of facilities for revision provided:

- **Spontaneous Revision of Topic**: After the completion of any topic, the spontaneous revision option is enabled and the subject can go back and browse through the conversations with the agent. Completing the flashcards associated with the topics is considered as the completion of revision.

- **Scheduled Revision of Topic**: Once the course is completed, in order to ensure that the effort spent during learning the course is not wasted, the conversational agent draws up a revision schedule.

The revision schedule consists of a five-term revision separated at certain intervals of days in order to implement Spaced Repetition. The revision schedule is based on mean average **retention score** for all the topics present in the course.

3.9 Retention Score

Retention score is calculated for each topic with the following equation:

$$R = 10e^{\frac{-t}{s}} \tag{2}$$

$$\text{where } R = \text{Retention Score,}$$
$$t = \text{time passed since topic or revision session completion}$$
$$s = \text{Retention Strength of the topic}$$

Figure 5: Retention Score Calculation

Retention Strength has a base strength. It is increased when scheduled revision is performed and also when the learner's notes are modified regarding the topic. Retention Strength depends on:

- Number of topic notes the user has added

- Number of revision notes the user has added

- Number of questions the user has proposed and answered

- Number of terms the user has defined

- Number of Equation Notes the user has defined.

Retention score has been created to put a quantitative figure on the subject's retention. Though there is no evidence for any correlation existing between the retention score and the actual retention, it acts as a proxy for putting a number to the decaying knowledge.

- When the user completes the topic, R = 10 with t = 0. As time passes the Retention score decreases while strength remains constant.

- When a person completes a scheduled revision, revision is reset to 10 and strength goes up. Though spontaneous revision resets the retention score to 10, it doesn't affect the retention strength.

The purpose of Strength is to add an incentive whose effect is visible to the user. The following figure demonstrates the visible effects of revision.
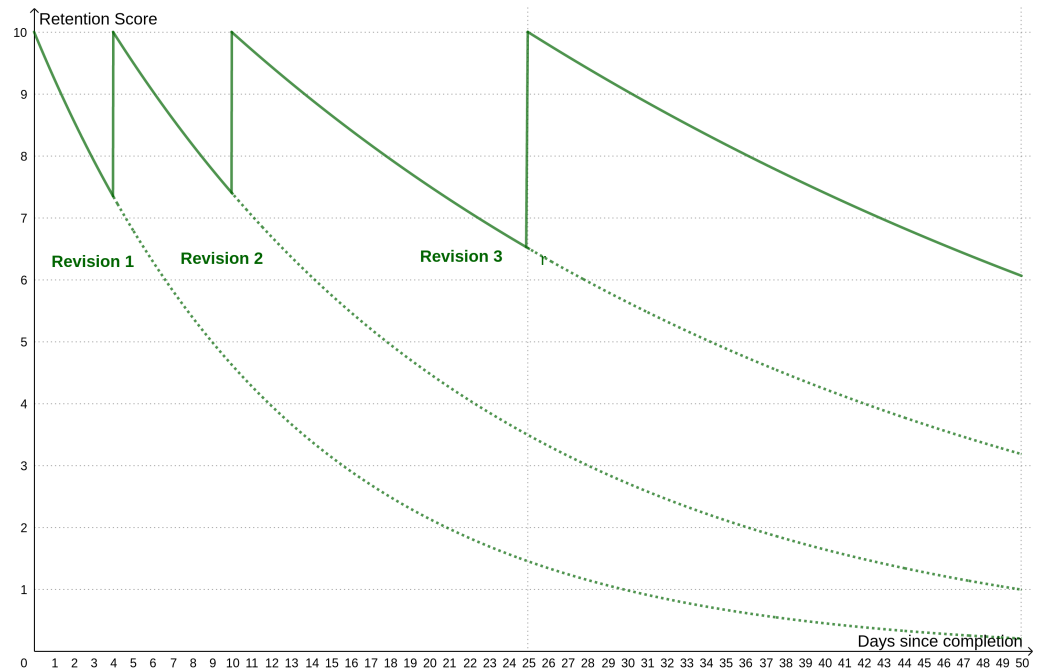

Figure 6: Retention Score plot against time passed since completion in days. Revision shown here are scheduled revision.

3.10 Incentives for completing the course and revision

The tutoring system takes into account that the learner will lack the discipline to complete through the course or to revise the contents frequently. In order to deal with the issue, the system incorporates incentives.

1. Incentives for course completion is that the personalized notebook and flash cards are generated only after course completion.

2. Incentives for revision is to keep up the retention score

3. Incentives for adding notes, terms, and questions is to increase the retention strength for the retention score.

The intelligent tutoring system outlined above is a basic framework for building a dialog based tutoring system. It has been designed to be extensible, customized to the needs of the

topic. Such a system will help anyone who wishes to learn difficult topics without a human actor.

4 Advantages of the tutoring system

The tutoring system is an effective tool as a standalone software as well as an aid to the educators.

1. **Extensible System** : The features of the tutoring system can be extended to accomodate features that help the learner learn.

2. **Faster completion**: Text based interactive courses on average takes less than 50% of the time to complete than video based courses.

3. **Persistent Attention Span**: The text based interactive courses help the learner to have greater attention span as the tutor doesn't proceed without the input of the learner.

4. **Complete Learning Framework**: The intelligent tutoring system helps learners learn, understand and retain.

5. **Journaling**: The interactions and conversations with the system serves for journaling purpose.

6. **Identifation of Knowledge Gaps**: The system allows the assessment of the learner's understanding at a low level using Learning Atom. This allows for learners to self evaluate the gaps they have in their knowledge and self-correct.

7. **Fast Iteration of Courses**: The course scripts can be improved upon feedback for the learners. As opposed to video based courses, the text based courses can be iterated rapidly.

5 Limitations of the system

1. **Tacit Knowledge**: Although, the framework can be used to convey explicit knowledge, tacit knowledge can be difficult to convey through the medium. It refers to the kind of knowledge that is difficult to transfer to another person by means of writing it down or verbalizing it. Examples can be learning to ride a bike etc.

2. **Cannot replace human**: A personalised dialog based assistant cannot match the effectiveness of an efficient human tutor. However, it can set it as a benchmark and can be improved to move towards that direction.

6 Future Scope

- **Addition of Natural Language Processing Abilities** The blueprint of the tutoring system demonstrated above shows that a dialog based agent doesn't require the abilities

to process natural language. However, incremental addition of such capabilities can be added to better facilitate learning and being more adaptive to the learner's need.

- **Building a social simulation game** Some tacit knowledge are learnt when the learner interacts with the real world. This cannot be communicated directly to the learner. A social simulation game which is sub-genre of the Life Simulation game can be an effective platform to learn such knowledge. Examples of such game include The Sims Series by EA, Animal Crossing etc. The game can be modified to mimic the real world with programming of actions and their resulting consequences. This game can be used by the tutoring system, Primer to boot up relevant scenario for the learner and acting as a guide in the game. For example, if the learner wishes to learn what does it take to become a politician, a game with the setting for the protagonist to be a politician can be booted up. This will enable the learner to learn through exploration and virtual experience.

7 Conclusion

We have proposed a tutoring system using a dialog based agent without the agent possessing capabilities of processing natural language. We started with the usual framework of an instructor model and converted it into scripts which can be used by the conversational agent to instruct learner. This was done by using Revised Bloom's Taxonomy to frame learning objectives and using principles of Known-to-Known Mapping and the Prior and Posterior Self-Learning. The system was extended using features such as automatic notebook generation, automatic flash cards generation and spaced repitition to enable the learner retain the course concepts for a longer time. Though, the system might be ineffective at the beginning, it can be improved by additions of new features and iterations of course scripts based on the Explainability metric.

8 Acknowledgements

9 Exhibits

References

[1] L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths, and M. C. Wittrock. A taxonomy for learning, teaching, and assessing: A revision of blooms taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman*, 2001.

[2] B. S. Bloom. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6):4–16, 1984. doi: 10.3102/0013189X013006004. URL https://doi.org/10.3102/0013189X013006004.

[3] T. M. E. BLOOMS. *Blooms taxonomy of educational objectives*. Longman, 1965.

[4] J. T. Gatto. *Dumbing Us Down: The Hidden Curriculum of Compulsory Schooling*. New Society Publishers, 2017.

[5] P. Gupta. Liberating education from the chains of imperialism. *Learning Societies: A Reflective and Generative Framework. Udaipur: ShikshantarThe People's Institute for Rethinking Education and Development*, 2000.

[6] D. R. Hofstadter. *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought.* Basic books, 1995.

[7] I. Illich. *Deschooling society*. Harmondsworth, Middlesex, 1973.

[8] P. Lockhart. A mathematicians lament: How school cheats us out of our most fascinating and imaginative art form. *New York, NY: Bellevue Literary Review (Quotation retrieved August 1, 2016, from http://www. goodreads. com/quotes/tag*, 2009.

[9] C. Newport. *Deep work: Rules for focused success in a distracted world*. Hachette UK, 2016.

[10] B. A. Oakley. *Mindshift: Break through obstacles to learning and discover your hidden potential*. Penguin, 2017.

[11] N. Stephenson. *The diamond age*. Penguin UK, 1998.

[12] J. Stockard, T. W. Wood, C. Coughlin, and C. R. Khoury. The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research*, 88(4):479–507, 2018. doi: 10.3102/0034654317751919. URL https://doi.org/10.3102/0034654317751919.