Original Paper

Design and Development of Synthetic Thought Generators for Books that are Important but Seldom Read

Sugata Mitra[1]

[1] Tataha Kim Laboratory, NIIT University, Neemrana, Rajasthan 301705, India

E-mail: sugata.mitra@niituniversity.in

Abstract

There are books that are central to human development and that have changed history over centuries. While we are aware of them, we know we will not read them. Our understanding of these influential ideas are from other people or secondary sources. Can a book respond like a person? We describe the design, development and preliminary testing of the "Synthor" (Synthetic Thought Generator) – a computational scheme that attempts to mimic conversation with a person who is knowledgeable about a book.

Keywords: Educational Technology, literature, reading, bots

1. Introduction

There are books that have changed the course of human history. However, it is ironic that while most of us are aware of these books, we know we will never read them. For example, Freud's "Interpretation of dreams" or Marx's "Das Kapital" have caused immense changes in the world we live in. We know this, but are unlikely to read these books (Wolf, 2018). There are too many of them. We prefer instead to knowing the influential ideas in these books from other people, who in turn may not have read them. The powerful ideas that have shaped humanity travel through generations much like a game of 'Chinese whispers', often resulting in distortion, ignorance, misunderstanding and even malice.

Ideas shape human civilization. Since the evolution of human sentience, transmission of ideas across generations have created the future for our species. Ideas are predominantly transmitted orally, or through writing. Written documents form the most important medium for the transmission and indeed the generation of new ideas. So important is this for human history that the period before writing was invented is referred to as 'pre-history'.

Due to the instant availability of information through the internet, people read less books. While this may not affect human productivity, perhaps even increase it, there are profound effects on children and society (Keller & Just, 2009). New ideas are generated through cross-fertilization and are dependent on a large pool of ideas to start with, as in most evolving systems. The education system tries to generate such a pool of ideas by encouraging learners to read as many books as possible. This method does not work well any more as learners' attention spans have reduced due the rapid access capabilities of digital technology. Moreover, learners are increasingly isolated as digital technology does not require physical presence for communication. Loneliness and low attention span can combine to create grave psychological imbalances (Wilson et al., 2013).

New ideas can be generated by gathering together people with a rich pool of ideas from books they have read or are their own. Indeed, exchange of ideas is the reason for academic conferences (Ridley, 2010). There can be little doubt that a conference on evolution with Charles Darwin present or a conference on governance with Plato in attendance, would be enriching. In this paper, we describe how digital technology can help create the semblance of such 'presence' and thereby increase the possibility of the production of new ideas while reducing the disadvantages due to isolation and lack of attention (Gunasekaran et al., 2016).

2. Design Considerations

We need to design a system that can present ideas or 'thoughts' of an author in response to statements, questions or comments from persons interacting with it (Yao, 2016). We will call such a system a "Synthetic Thought Generator", abbreviated to SYNTHOR.

A book is a collection of sentences that need to be read in sequence. This, indeed, is the main reason why people who are short of time, or with a deficit of attention, cannot read books, even if they want to. As discussed above, the company of a person who has read a book can partially compensate for someone who has not read the book. It is important to point out that reading a few lines at random from a book cannot be a replacement for the company of someone who has read the entire book. Similarly, looking up keywords from the index of a book will not provide a similar replacement, unless the book in question is a manual or a collection of indexed information. What then is the difference between the company of a person who has read a book as against other mechanical means of finding a relevant portion of a book?

I suggest that the answer to this question lies in the stochastic nature of human thinking. The response of a human being to a statement, question or comment is constructed out of a 'cloud' of possible responses, the construction being dependent on prior knowledge, memory, recall and mood, among many other psychological or physical factors. As a result, the response to the same verbal stimulus may be different at different times (see for instance, Catalanotto, 2016). This variety is what attracts people to each other instead of to a mechanical algorithm that always produces the same response to a particular stimulus. If conversation with a Synthor is to have even a semblance of the usefulness of a conversation with a person, its design must incorporate uncertainty.

3. Design of a Synthor

We start by defining a mind as a collection of thoughts, each thought being represented as one or more sentences. This definition is a purely functional one for the purpose of creating a Synthor and does not, necessarily, have any connection with what an actual mind is – if there is something called an 'actual mind'.

Each sentence in the cloud of sentences that constitute a Synthor's mind has an equal probability of becoming a response to a stimulus such as a statement, a comment or a question. However, context, keywords, and appropriateness (mood) determine the sentence, or sentences, that will be actually selected for response. In other words the probability of sentences appearing as a response are adjusted until a sentence or set of sentences 'collapse' into an actual response. This is analogous to the way the state of a fundamental particle in nature is revealed as the probabilities of many possible states collapse into one particular state in response to a measurement (see for instance, Wolchover, 2020).

The design of a Synthor will, therefore, consist of methods for determining the probable context for a conversation, extracting keywords that represent this context, then adjusting the probability of selection using a 'mood' that will represent the appropriateness of a response, and, finally, presenting the most probable response from a final list of high probability responses. In addition, each step in this selection process can be altered by the probability of forgetting to adjust a value.

A Synthor with a mind consisting of the contents of just one book would respond as a whimsical but knowledgeable reader of that book.

4. Development and Preliminary Testing of Synthors

The first Synthor developed was a desktop computer application using the above design principles. It consisted of a cloud of about 17,000 sentences, created from all publications of the author. Five probabilities were used to adjust the context, keywords, mood, forgetfulness and appropriateness of sentences from the cloud before presenting the most likely sentence as a response to a stimulus from a user of the system. The interface for the system was simple – the user typed into a box in one part of the screen and the Synthor responded by typing into another box on the screen. Response times were in fractions of seconds and resembled a rapidly typed text exchange on the internet. The five probabilities were adjusted until the author (who is the developer of the system) felt that the responses were adequate for testing. The final system was called "First Synthor". Box1 shows an example of a

conversation:

Box1: Part of a conversation with the First Synthor.

> **User:** I was trying to publish this software
> **First Synthor:** ….While the hardware and media for traditional and multimedia publishing are converging, there is a large variety of software tools that are making this convergence possible.
> **User:** but it is not easy
> **First Synthor:** Identification is done by talking to the children, studying their level of interest and curiosity, and more, selecting children who lived around the kiosk (for easy accessibility).
> **User:** not easy to convert to a windows app
> **First Synthor:** Not only that, you cannot run the Windows version on smartphones and the Android App on a Windows or Apple computer.
> **User:** absolutely!
> **First Synthor:** Turning the 'minds eye' towards itself is easy and it immediately reveals the Self as Abstract Consciousness.

Four users interacted with the First Synthor in four sessions lasting for about 15 minutes each. Three of these users were people who were familiar with most aspects of the authors publications while the fourth user (referred to as X) was unaware of any of these publications.

Reactions from all users after their sessions contained words such as, "It's incredible", "It is quite insane", "Scary", "It changed the topic", "It was in a bad mood" and "wish it would tell me more". After the last session, X expressed a desire to discuss a certain aspect of children's education that the author had published sometime ago.

Using the design and experiences described above, three more Synthors were developed. A Synthor called "Darwin Synthor" was developed using the text from Charles Darwin's book, "On the Origin of Species". Next, a "Kant Synthor" was developed using the text from Immanuel Kant's book, "Critique of Pure Reason". These books are generally considered among the most influential of all time. Finally, we created a Synthor from the text of the Rig Veda, called the "RigVed Synthor". The Rig Veda consists of over 10,000 verses, composed, and transmitted orally since, possibly, the end of the last ice age, and finally put into writing around 1200 BCE (Cairae, 2014). It was created by many persons over the ages and forms one of the oldest publications of humanity. As a result of its antiquity, its context and contents are difficult, if not impossible, to comprehend and few people have read it in its entirety, although it has affected the languages and cultures of most of Asia and the rest of the world.

It is important to state that creating new Synthors is not merely a substitution of one text database by another. For each text cloud, the correct representation of ideas as groups of sentences is a vital task. Even more delicate and difficult is the task of adjusting the five probabilities that define the behavior of each Synthor. In the author's experience, the creation of a Synthor is closer to sculpting than to programming – as bits of code need to be moved and shaped according to the nature of the Synthor being created.

Each of the Synthors described exhibit the character of the book they are created from. This reflects in their language and the length of their responses. The relationship between a response and the stimulus text is sometimes not evident and this lack of apparent coherence seems dependent on the nature of text that forms the mind of a Synthor. Box2 shows an example.

Box2: Responses of four Synthors to the same question.

> **Question: What role do our senses play in what we know?**

> **First Synthor:** If the schools played a more proactive role in allowing children to use the PLS during school hours, the data suggests that the impact on children would be larger.

> **Darwin Synthor:** The possibility, or even probability, of inherited variations of instinct in a state of nature will be strengthened by briefly considering a few cases under domestication. We shall thus also be enabled to see the respective parts which habit and the selection of so-called accidental variations have played in modifying the mental qualities of our domestic animals.

Domestic instincts, as they may be called, are certainly far less fixed or invariable than natural instincts; but they have been acted on by far less rigorous selection, and have been transmitted for an incomparably shorter period, under less fixed conditions of life.

Kant Synthor: ….bodies are not objects in themselves that are present to us, but rather a mere appearance of who knows what unknown object; that motion is not the effect of this unknown cause, but merely the appearance of its influence on our senses; that consequently neither of these is something outside us, but both are merely representations in us, hence that it is not the motion of matter that causes representations in us, but that motion itself (hence also the matter that makes itself knowable through it) is a mere representation; ….

RigVed Synthor: Rig Ved - Mandalam 8 - Shloke 13 HYMN XIII. Indra.

29   These his victorious followers bold in the heavens the place they love, Leagued in the heart of sacrifice, as well we know.

30   That we may long behold the light, what time the ordered rite proceeds, He duly measures, as he views, the sacrifice.

5. Questions

Synthors can be developed to contain the text clouds from more that one book. The behavior of such Synthors are likely to be different from Synthors based on text from single books. Would these appear "wiser"?

One of our biggest challenges is about the use of text from works of fiction for the creation of Synthors. For obvious reasons, sentences from works of fiction are highly contextual and depend on the readers understanding of imaginary places, people, things, and sequence of events. How are ideas from fiction internalized? Can such a process be integrated into software?

In this paper we are able to report the design and development of Synthors. However, more research is required to find out what effect such Synthors have on people, particularly learners. Are these effects at all similar to being surrounded by the authors of these monumental works? If so, Synthors could catalyze the development of new ideas in the future.

References

Cairae, H. M. (2014). *An Aryan Journey*. Rupa Publications.

Catalanotto, Daniele. (2016). *It's now time to design context aware websites*. https://service-design.co/its-now-time-to-design-context-aware-websites-43546071aa34

Gunasekaran, S. S., Mostafa, S. A., Ahmad, M. S., & Tang, A. (2016). Cross-fertilization of ideas in Collective Intelligence Model (CIM). 2016 2nd International Symposium on Agent, Multi-Agent Systems and Robotics (ISAMSR), *2016*, 27-33.

Ridley, Matt. (2010). https://www.wired.co.uk/article/ideas-having-sex-matt-ridley-steven-berlin-johnson

Robert, S., Wilson, Patricia A., Boyle, Lei Yu, Lisa L. Barnes, Julie A. Schneider, David A. Bennett. (2013). *Neurology*, *81*(4), 314-321.

Timothy A. Keller, Marcel Adam Just. (2009). Altering Cortical Connectivity: Remediation-Induced Changes in the White Matter of Poor Readers. *Neuron*, *64*(5), 624-631.

Wolchover, Natalie. (2020). *What Is a Particle*? Quanta Magazine. https://www.quantamagazine.org/what-is-a-particle-20201112/#back1

Wolf, Marianne. (2018). *Skim reading is the new normal*. The effect on society is profound. https://www.theguardian.com/commentisfree/2018/aug/25/skim-reading-new-normal-maryanne-wolf

Yao, Maria. (2016). *Chatbot UX – Does Conversation Hurt Or Help*? https://www.smashingmagazine.com/2016/11/does-conversation-hurt-or-help-the-chatbot-ux/