A Knowledge Ecosystem - Deep Learning and Education

Fanli Zheng (Christian Ramsey) & Haohan Wang

2018-11-14

Contents

Preface **6**

1 Introduction **7**

2 Primary Concerns **11**
- 2.1 Passive Consumption and Untested Knowledge 11
- 2.2 Knowledge Ecosystem Example 11
- 2.3 The Problem of Knowledge Representation 13

3 Concepts **15**
- 3.1 Knowledge Footprint 15
- 3.2 Knowledge Journeys 16
- 3.3 Collective Human Knowledge Graph 17
- 3.4 Knowledge Ecosystem Example 18

4 Implementing the Knowledge Ecosystem **19**
- 4.1 Problem Formulation 19
- 4.2 LEARNING + FEEDBACK 20

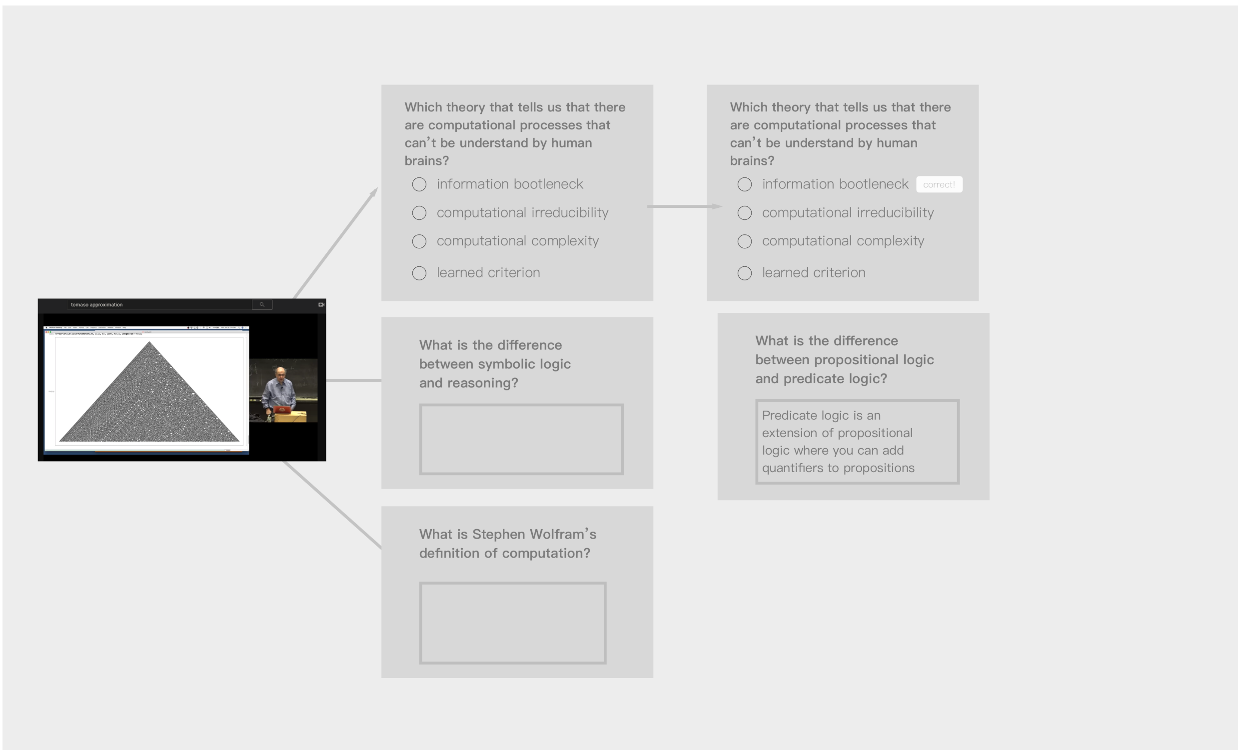
5 KNOWLEDGE GRAPH **31**
- 5.1 KNOWLEDGE JOURNEYS 31
- 5.2 Data and Annotation 32

6 Conclusion **35**

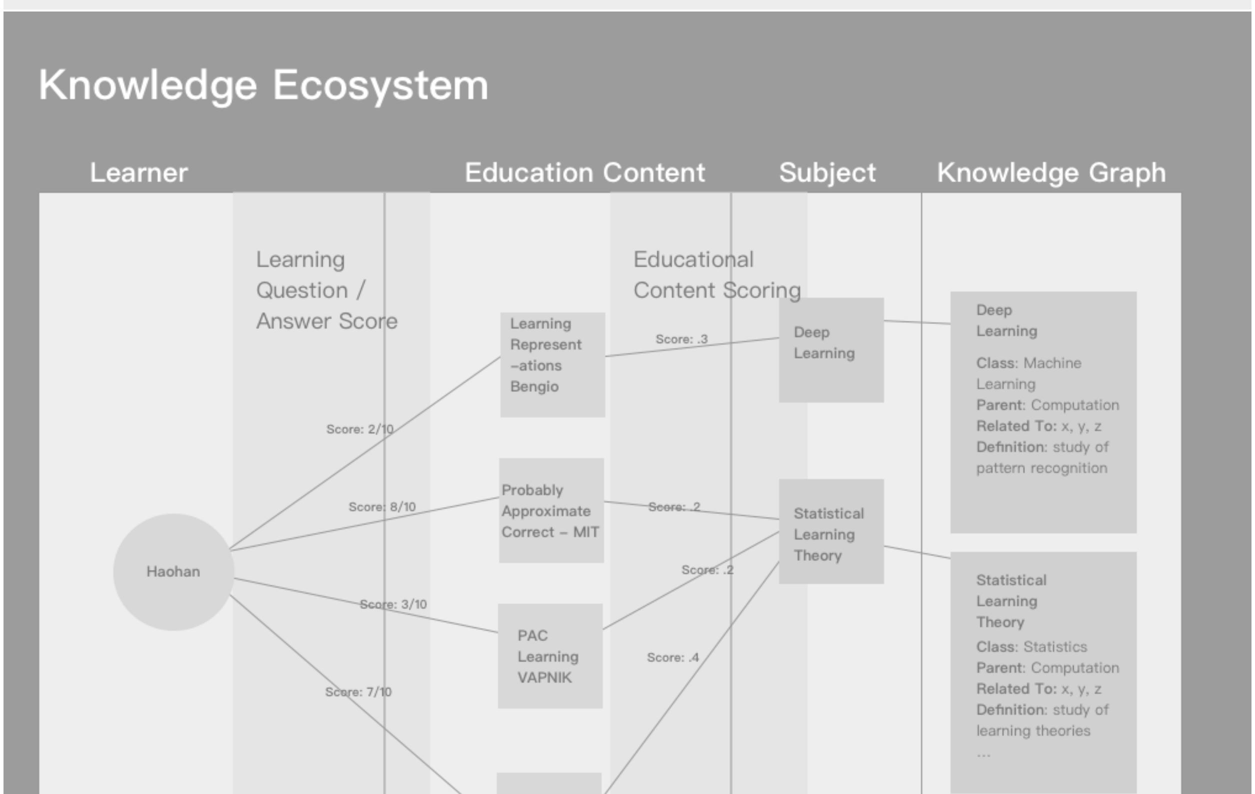
7 About Authors **37**
- 7.1 Haohan Wang 37
- 7.2 Fanli (Christian) Zheng 37

8 Links **39**

Preface



Which theory that tells us that there are computational processes that can't be understand by human brains?
information bootleneck
computational irreducibility
computational complexity
learned criterion
Which theory that tells us that there are computational processes that can't be understand by human brains?
information bootleneck
correct!
computational irreducibility
computational complexity
learned criterion
What is the difference between symbolic logic and reasoning?
What is the difference between propositional logic and predicate logic?
Predicate logic is an extension of propositional logic where you can add quantifiers to propositions
What is Stephen Wolfram's definition of computation?



Knowledge Ecosystem
Learner
Education Content
Subject
Knowledge Graph
Learning Question / Answer Score
Educational Content Scoring
Learning Represent -ations Bengio
Score: .3
Deep Learning
Deep Learning
Class: Machine Learning
Parent: Computation
Related To: x, y, z
Definition: study of pattern recognition
Score: 2/10
Score: 8/10
Probably Approximate Correct – MIT
Score: .2
Statistical Learning Theory
Haohan
Score: .2
Score: 3/10
PAC Learning VAPNIK
Score: .4
Statistical Learning Theory
Class: Statistics
Parent: Computation
Related To: x, y, z
Definition: study of learning theories
...
Score: 7/10

Chapter 1

Introduction

Let's start with a future view of an individual's education. Many of us have used the internet to educate ourselves with the abundance of medium to high quality videos, papers, articles, podcasts and how-tos being uploaded from numerous individuals, groups, and institutions like never before (60 hours of video are uploaded to youtube.com every minute).

Let us imagine that all of what you have learned online, throughout the entirety of your life, from the hundreds of Youtube videos, Wikipedia articles, Nature papers, and podcasts you've read, watched, or listened to, were all added structurally to your **knowledge journey**, and what if that journey could be consolidated into what we might call a **knowledge footprint** that could be shared with others? Could this replace static degrees? Or augment them to be more inclusive of a learner's true knowledge?

Our current approach to education is to treat knowledge acquisition like a chapter in the individual's life that is limited to one or more formal places. This is misleading since we accrue knowledge from everywhere and most recently the internet has become a primary source of knowledge acquisition but has gone mostly unaccounted for in terms of recognition (i.e. watching a whole series of Youtube lectures on the Neurobiology of Depression or Discrete Mathematics goes mostly unnoticed when someone views one's resume or by simply looking at their degree.). This makes it much harder for people to switch to working and exploring outside of their degree area. Knowing rigourous mathematics and not having a degree in it, is said to be surprising, therefore the current "thumbnail view" of an individual's knowlede is neccessarily inadequate to the new mediums of knowledge acquisition.

The ideas behind this *knowledge ecosystem*, presents only one of many possible solutions to bringing our education system into modernity. The goal of it would be to promote the long held idea of the life-long learner. Moving away from the "education chapter"" of an individual's life to the individual as an evolving learner; learning the necessary skills for what life presents them with today or might tomorrow. It would (combined with traditional education) show us a more accurate depiction of a learner's knowledge and theorefore that of a society's collective knowledge.

Visualised over time, we could begin to capture a learner's so called **knowledge journey**. Composed of every piece of content they've gained knowledge from mapped to the *human knowledge graph*. Showing how an individual has traversed through the world of human knowledge.

This would also serve as a way for others, who may be on a similar **knowledge journey** to connect with their cohort. This could be the start of meetups, study groups, and so on.

For those who are looking for a change, they may find different journeys that help them decide what step to take next. You would also be able to connect someone's occupation to their **knowledge journey**.

On aggregate, we could begin to cluster similar **knowledge journeys** through unsupervised learning, which might lead to completely new journeys or combinations of subjects at differnet levels that others may be inspired to follow.

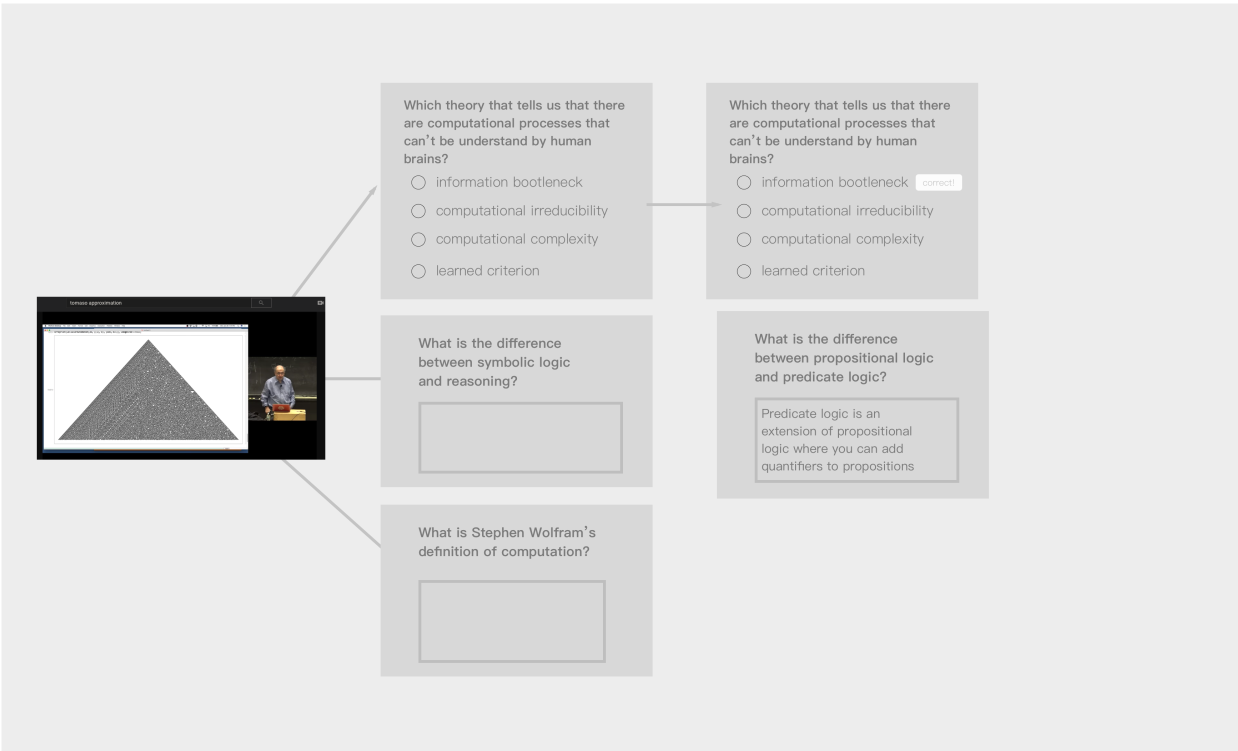


Figure 1.1: Image title

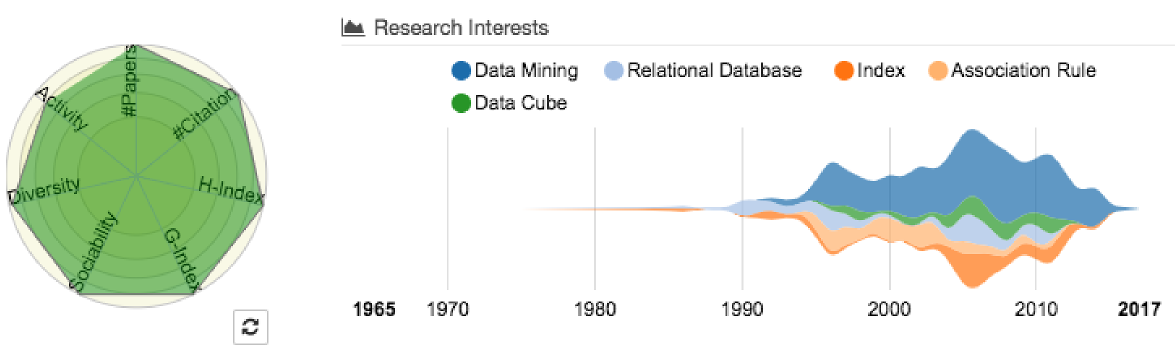


Figure 1.2: Example at aminer

In this essay, I will propose a **knowledge ecosystem**, a new way of approaching education that attempts to build a more accurate depiction of a learner's true knowledge. It will require significant effort to bring to life but I believe the benefits will outweigh the costs. I will talk about how we can use machine learning, deep learning in particular, to help create and support a **knowledge ecosystem** which is made up of a **knowledge footprint**, **knowledge journeys**, and a **collective human knowledge graph**. We will also introduce current advances in deep learning that would enable us to take the space of unstructured educational content on the web and do the following, - classify content to higher level subjects - map content unto the human knowledge graph - test a learner's knowledge of recently viewed educationcal content through questions and answers, no what matter the subject.

I will also argue that this imagined future is not only **desirable** for society but something similar is required to insure individual's knowledge are well represented in a time where the pace of change is rapidly speeding up. Let us not forget, that even software engineering is currently being recreated with machine learning as a key pillar which wasn't much of a thought 5-10 years ago.

As the future pushes us forward, it is tantamount that we have adaptative systems that can represent our current knowledge and also make us predictable to others given the future pushes us to knowing more than ever and knowing who to collaborate with to apply such knowledge.

This hypothetical future isn't just conceptual, most of what I will present to you today is currently feasible due to the most recent advances in machine learning, and in particular deep learning

In the last section of this essay I will review what has been proposed and also call other researchers, teachers, and designers to collaborate on such an ecosystem, even if it is just in part.

*For the purpose of this essay we will talk mostly about digital knowledge acquisition and leave the reader to extend the basics to knowledge obtained elsewhere.

Chapter 2

Primary Concerns

There are 3 popular concerns that I will attempt to address in this article about online knowledge acquisition that stand in the way of having an adaptive and reliable knowledge ecosystem. I will attempt to present a system that can sufficiently overcome each of the concerns here and in the implementation section.

There are as follows:

- **Passive Consumption** - most of online content is viewed passively by the learner and the result of passive consumption is that learner's do not grasp the concepts or master the content being taught.
- **Untested Knowledge** - even if the learner was engaged while viewing a piece of educational content their knowledge is untested and therefore it isn't clear if they've mastered the content accurately and in some sense holistically.
- **Knowledge Representation** - even if the learner was engaged (1) and their knowledge was tested (2), simply knowing the counts or types of video they watched doesn't make their knowledge predictable and useful to others. In fact, even the learner may be unaware of all of what they've viewed.

2.1 Passive Consumption and Untested Knowledge

How would such an ecosystem insure us against passive consumption?

Scenario #1: A learner goes online and begins watching a series on Machine Learning. How do we engage and test a user's knowledge?

Proposition: Using advances in deep learning, we propose a dual question and answer generation given the educational content.

Result: A learner gets a set of questions and multiple choice answers throughout the video. Keeping the user engaged and sharp to ensure they can answer each of the questions.

As you can see, I've bundled passive consumption and untested knowledge because our proposed ecosystem approaches both of these by always testing knowledge. I will show the current results in deep learning in the implementation stage.

2.2 Knowledge Ecosystem Example

Given a piece of educational content, our knowledge system will generate a set of questions and answers that theoretically capture the major concepts and facts that the learner should know after viewing a part or the content in whole.

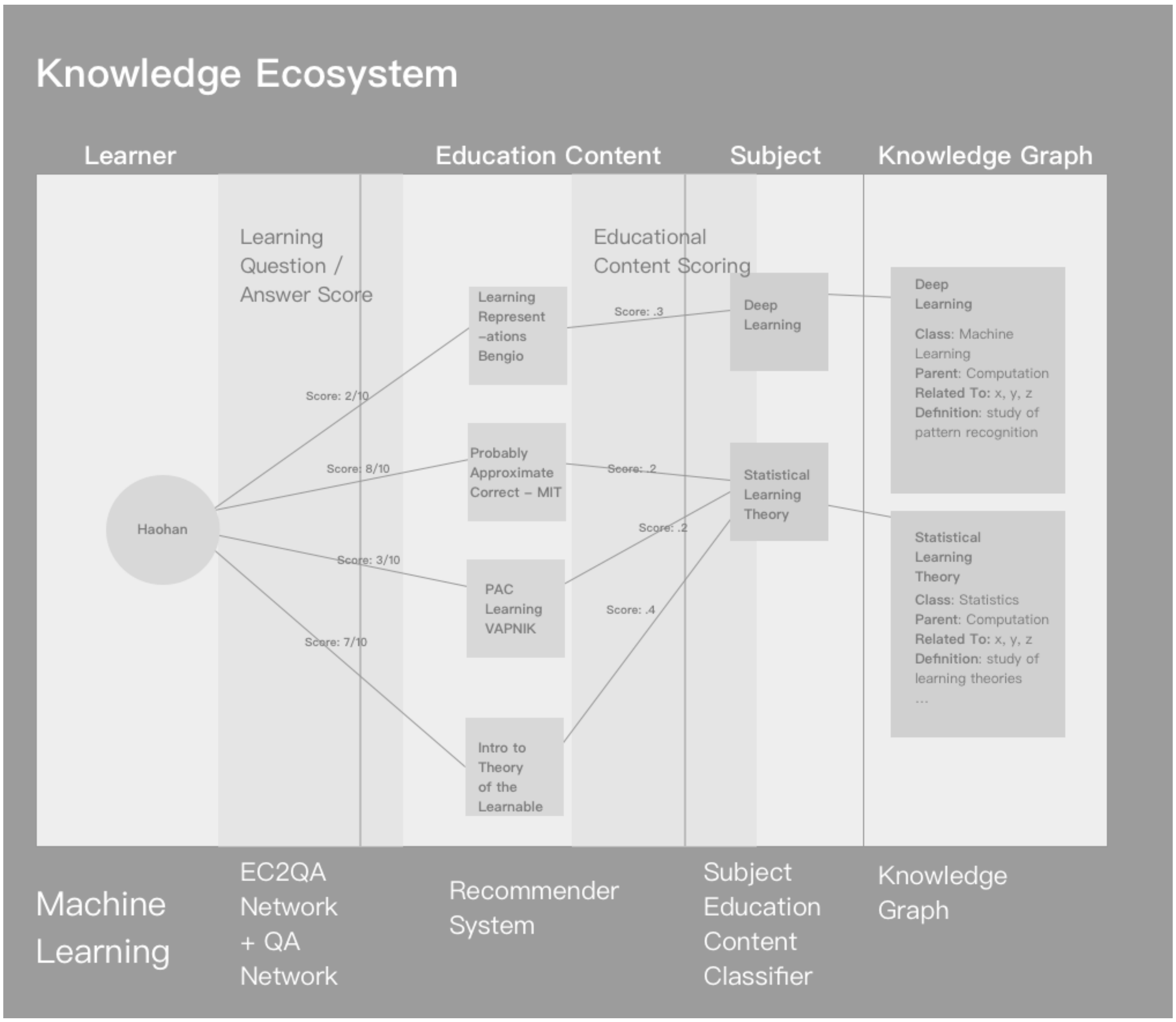


Figure 2.1: Knowledge Ecosystem

You can imagine watching a Youtube video and after a learner views 15 minutes of an hour long lecture on computational complexity a quiz is presented (a set of questions and answers conditioned on the past 15 minutes of video,) and the results are recorded. In the future we would also be able to use the knowledge graph

The knowledge system would only consider content that has been watched with some engagement or if they can test out of the content.

2.3 The Problem of Knowledge Representation

> Given how many learner's online knowledge acquisition varies and has been invisible up to now, how we can best represent their knowledge?

Scenario #2: A learner has a degree in Public Health, but since graduating, they've been studying machine learning for the last 3 years. The learner now wants to apply to a job that requires Health and Statistics. How do we represent their traditional and updated knowledge?

This is a tricky problem that goes beyond any given algorithm. The exact design of a **knowledge footprint** and a **knowledge journey** has been attempted and we will not cover that in depth here. The proposed system presupposed the design of the knowledge footprint.

There are two problems, > how do we reduce someone's knowledge (in this case a set of educational content and their respective scores) into a symbol

Proposition: We introduce **knowledge journeys** and the **knowledge graph** as a way to make sense and structure a learner's knowledge acquisition. The collective **knowledge graph** will tell us about the subject the learner is studying and we can use this to compare to others and create a relative comparison.

Result: Reducing a learner's **knowledge journey** into a common set of elements that makeup into their **knowledge footprint** which would look similar to those with similar journeys.

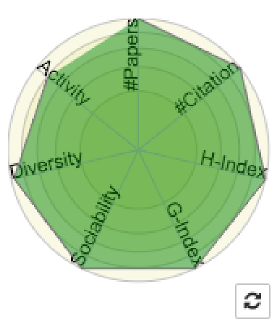
So the employeer, now familiar with the footprints can see the overlap between the current employee's and a prospective employee.

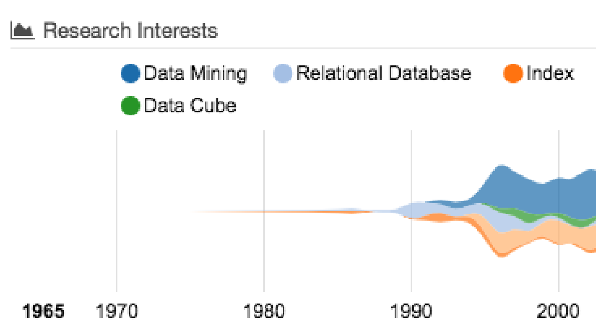
Chapter 3

Concepts

3.1 Knowledge Footprint

The concept of a knowledge footprint is a custom symbol or badge with a profile that represents one's education relative to that of others. This footprint should represent all of one's education (currently focused on digital) while balancing distinction and commonality with others.





This example from aminer is very helpful.

3.2 Knowledge Journeys

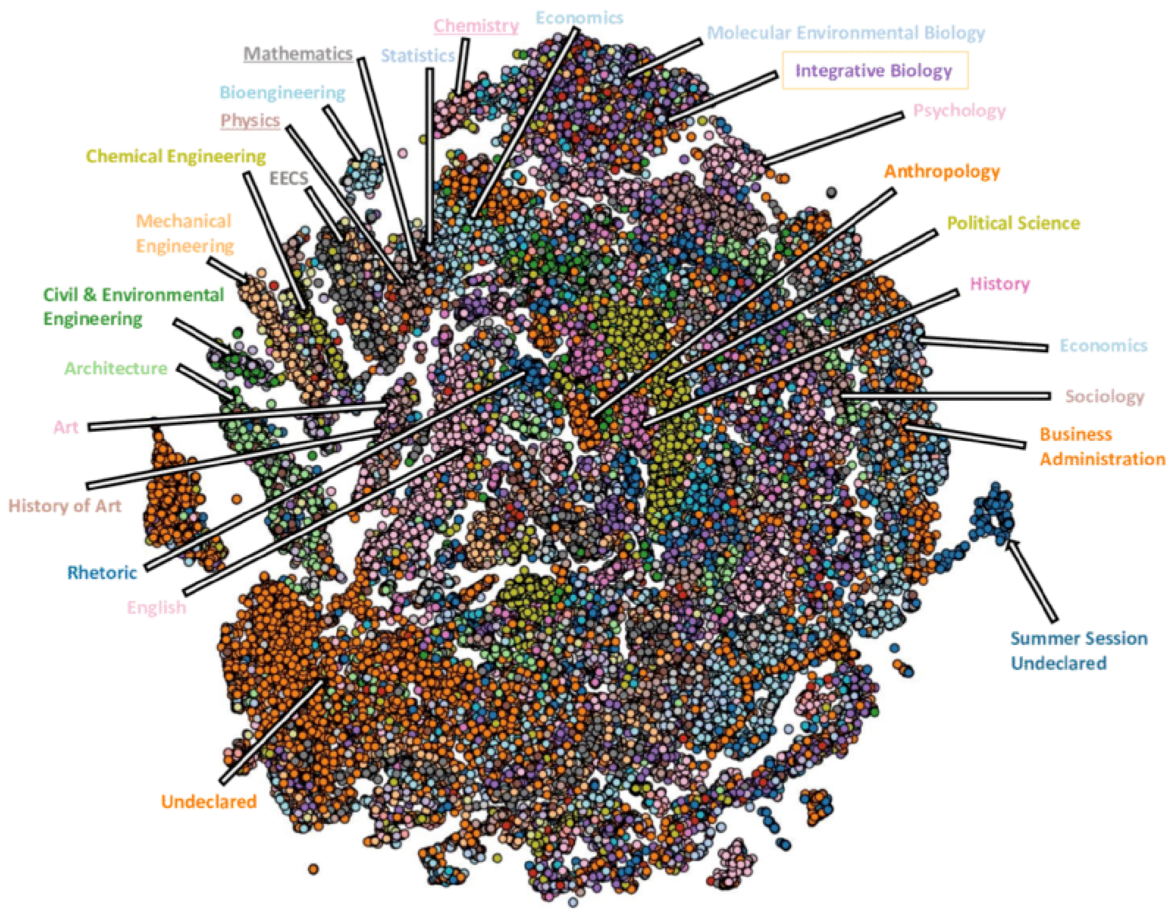
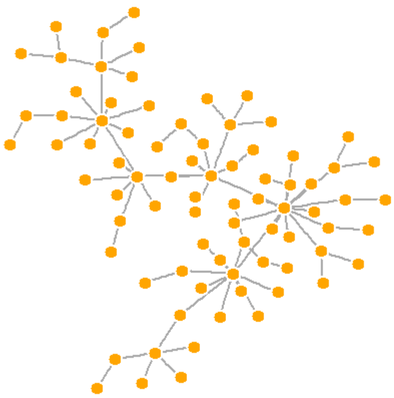


Image src: https://www.researchgate.net/figure/t-SNE-projection-of-the-embedding-of-all-learners-in-the-dataset-Major-lal
fig1_323391033

A **knowledge journey** is a somewhat simplified view of all of the educational content a learner has acquired over time. The journey should be a temporal representation of all of the subjects that one has viewed and been tested on. Knowledge journey's should be simple enough to compare but complex enough that the individual can go back to a particular moment in time and rewatch educationcal content they've viewed before.

3.3 Collective Human Knowledge Graph

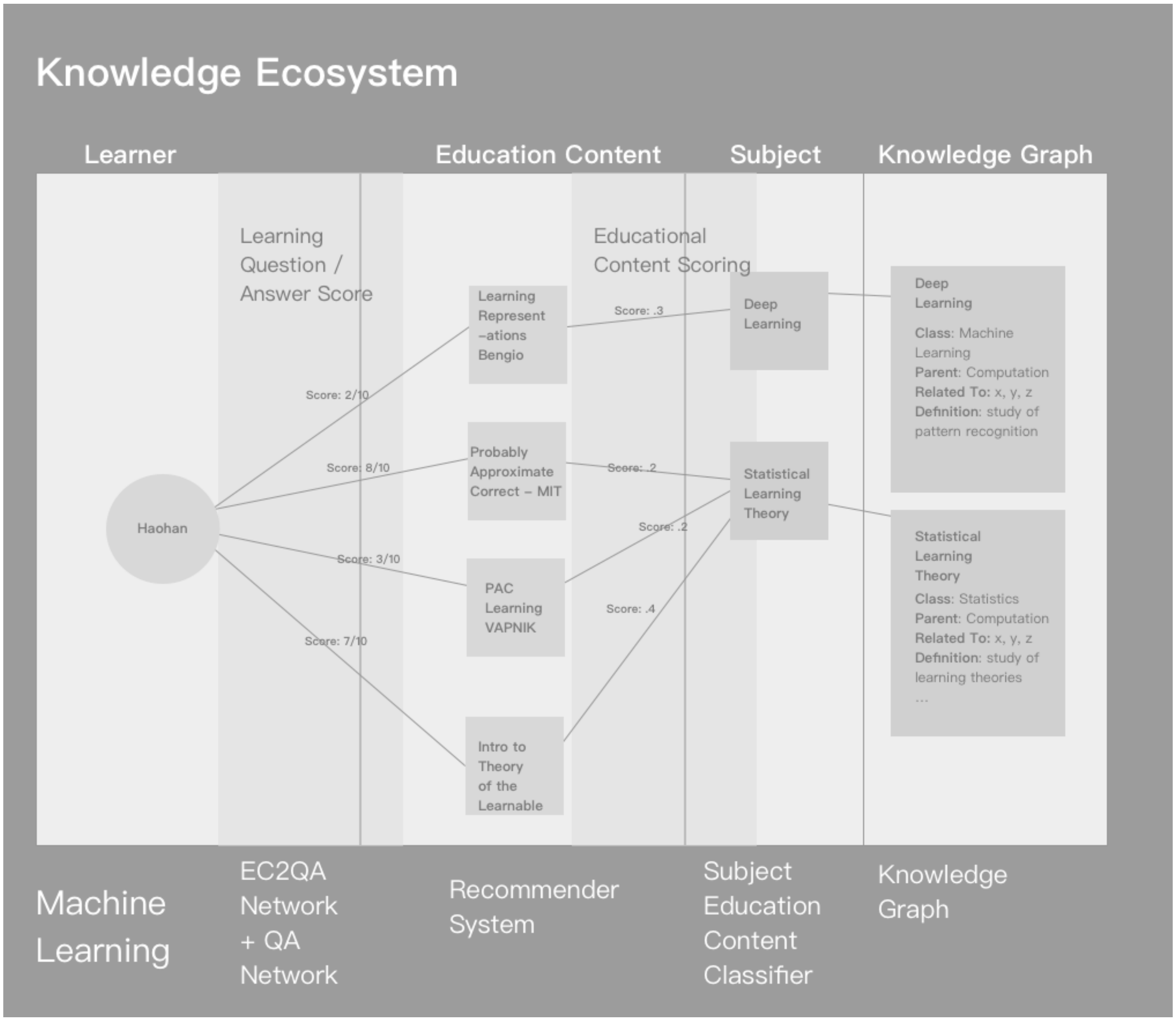


The collective human knowledge graph can be compared to Google's Search Knowledge Graph which points unstructured information towards structure. The graph should have all existing subjects that we are currently aware of (i.e. Abstract Mathematics, Discrete Mathematics, Statistics, Art, Sociology). Since each piece of educational content will be classified into a subject, all subjects will exist in detail within the knowledge graph.

3.3.1 EC2QA Network

We mentioned the EC2QA network earlier because currently we would have to cobble together multiple networks to make this work. Instead we will introduce a novel network architecture, EC2QA, to solve the problem of generating a set of questions and answer pairs for any given educational content (text, video, image, pdf).

3.4 Knowledge Ecosystem Example



Now that we are aware of each of the elements, let's talk about how they work in practice.

A learner watches a video titled 'Depression' by Robert Sapolsky. - The video is classified by a neural network as [Neuroscience, Mental Health, Psychology] - The subjects are then mapped to the **knowledge graph** which gives us more information about each subject - Using EC2QA or similar, a set of questions and answers are generated for every 15* minutes of video A learner is present with 5 questions to answer and scores 4/5 (80%) - The link between user and video takes up the score At the end of the video, the learner records a video summary and is evaluated with a 7/10 (70%) - The evaluation network looks at the **semantic and conceptual mututal information** shared between the original content and the learner's video summary - All scores are mapped to the video and also counted at the subject level A learner looks at their knowledge footprint - All scores should be calculated against all subjects coming from the knowledge graph and compared against other learners to generate the footprint

Chapter 4

Implementing the Knowledge Ecosystem

In this section, we are set to solve the following question:

> How might we approach designing such a system.

I will walk you through some possible implementations of the proposed knowledge ecosystem. I will be presenting the current research in machine learning that is relevant to each component of the knowledge ecosystem and also discuss this new artificial neural network architecture EC2VQA and one possible instansitation of that.

Note that this is not the blueprint, each of these components can be developed independently and vary from what you find here. This is a provocation for getting started on the knowledge ecosystem today.

4.1 Problem Formulation

Building such a digital knowledge ecosystem like the one we just visioned is not a trivial task. We would need to break this ecosystem into few key building blocks as below so that we can set out to tackle them one by one.

1. **LEARNING + FEEDBACK** – given a learner consumes a piece of educational content, reliably evaluate their knowledge and provide the feedback for improvement to support their learning. (credibility, rigour)
2. **KNOWLEDGE GRAPH** – general knowledge blueprint as a map to piece together all the content that is currently available. (relatibility, predictability)
3. **KNOWLEDGE JOURNEYS** – given a piece of educational content, classify it within a knowledge graph; given multiple learner journeys, create a way to customize their own growth journey while offereing a way for them them to compare, connect with and follow, other's journey (compare, traverse, curiosity)
4. **KNOWLEDGE FOOTPRINT** – given a learner's journey, collapse it into a representative symbol(s) (digital education footprint) (relatibility, stable but evolving system)

Much of the element #1 and #2 have been made possible with the recent breakthoughs in machine learning especially in the filed of deep learning. A few other pieces like the element #3 and #4 may require a more advanced framework that has not been proposed yet to best resolve. Let's first take a look at some methods that might come handy when applying to our problems.

4.1.1 Before we start...

As what we discussed above, one of the possible and optimal solutions so far for building our ecosystem is using current state-of-art deep learning algorithms in the related domains. Here I will walk you through some thought provoking research summary in a form of survey.

To best illustrate the problem and possible solutions, we will for now reduce the complexity of the problem. Keep in mind that our ultimate goal is to be able to apply our system to any type of online available educational content. As what we have mentioned earlier, educational videos seem to be one of the top options for people when it comes to knowledge acquisition. Let's explore some of the solutions that can enable us to apply our system on those video educational content.

4.2 LEARNING + FEEDBACK

4.2.1 Question Formulation

We consider learning + feedback as a key building block of the ecosystem which is the cure for our primary concern over passive knowledge consumption and untested knowledge of learners. Here, we will be addressing this problem and its possible solutions in deatails so that we can provide our learners a credible and interactive learning + feedback experience system to best support their knowledge acuqisition.

A salient question that we have to ask ourselves before deisgning such a system is that:

> "how can we take a piece of educational content and properly test a learners' knowledge and provide insightful feedbacks to support their learning?".

In previous years, deep learning research has taken up a similar problem titled Question Generation (QG) and Question Answering (QA).

Question Generation (QG) is part of NLP. The goal of QG is to generate questions according to some given information. It could be used in many different scenarios i.e. generating questions for reading comprehension, generating data from large scale question-answering pairs or even generating questions from images. Earlier approaches to QG mainly used human-crafted rules and patterns to transform a descriptive sentence to a related question. Recent neural network-based approaches represent the state-of-art of most of those tasks and this approach has been successfully applied to many other NLP tasks i.e. neural machine translation, summarization, etc. As the training optimization studies progress, the stability and performance improvements are gauranteed.

As for Question Answering (QA) task, it is a well-researched problem in NLP as well. Recently, QA has also been used to develop dialog systems and chatbots designed to simulate human conversation. Traditionally, most of the research used a pipeline of conventional linguistically-based NLP techniques i.e. parsing, part-of-speech tagging and coreference resolution. However, with recent developments in deep learning, neural network models have shown promise for QA. Further improvement i.e.attention mechanism and memory networks allow the network to focus on the most revevant facts such that they achieved state-of-art performance for QA.

Now we have some basic understanding of these 2 problems that we will be investigaing more in depth later. Consider the next question:

> "what types of questions & answers would be best to test a learner's knowledge given a piece of educational content (i.e. a lecture video)"

Let's say a learner is watching a video about hypothesis testing, after showing an example and given the data needed to test the hypothesis. Then the system will pause the video and ask:

1. What is the p-value for this test? (and provide multiple choices for learner to choose from)
2. Is this a 1-sided test? (answers provided would be: YES or NO)

3. How would you intepret the p-value in the context of this example.
4. Based on your calculation, summarize the conclusion of this hypothesis testing for me.

It is also ideal to ask learner the question as follows after showing the solution and explanation:

5. Tell me what you have learned through this example.

As shown above, we would call questions #1 and #2 close-ended questions; question #3 and #4 specific open-ended questions; and question #5 a general open-ended question.

So the formulation for both of these is to:

1. Generate close-ended question + Answers pairs
2. Generate specific open-ended question + Answers pairs
3. Evaluate and comment on the general open-ended answers

In terms of the close-ended, the answers can be well defined and evaluated. However, the process might be a little bit tricky when it comes to the open-ended questions. We will approach each of them here from the current research.

** Why deep learning? **

As we stated above, deep learning has achieved state-of-art performance in both QG and QA tasks. Now let's take a close look at why.

If you pay close attention to the question generation and answer generation set of problems, we can easily frame this problem into a general machine learning problem in which we need to find the relationship between the educational content and the meaningful question & answer pairs that is associated with the content. In other words, our problem could be simplified as learning a function that is capable of capture the relationship between our input and output, that is to say, appropriately map the educational content to the desired question and answer pairs with this function.

To the best of our knowledge, deep learning is one of the most optimal techniques currently developed to solve such an issue.

As we all know that machine learninag is a set of algorithms that can be used to parse data, learn from the data, and then apply what they have learned to make intelligent decisions. Or more specifically,

Deep learning is a subset of machine learning and it belongs to a family of representation learning. Inside this family, deep learning is good at sampling the features and have additonal layers for more abstract feature learning which will support the final goal of mapping the feature to the output.

Because of the above advantages, deep learning is known as one of the most flexible machine learning algorithms that can learn and map the **deep representation** from the data. Also, deep nueral networks architecture can be composed into a single differentiable function and trained end-to-end until it converges. As a result, they can help identify the suitable *inductive biases* catered to the training data.

Moreover, deep learning outperforms other techniques when the training data size is large which fits our siutation well. We have a large amound of educational content that is currently available on the web.

The large amount content creates another problem that can be avoided with deep learning as well, which is it's going to very troublesome to do feature engineering manually. When there is lack of domain understanding for feature introspection, deep learning is preferable.

In the end, deep learning really shines when it comes to many complex reserach problems such as NLP, Visual Recognition and Speech recognition. For solving our task, all those domians will possibly be invloved.

4.2.2 Question Generation

First, let's take a look at question generation (QG) problem.

The ideal goal of an automatic question generation is to generate a question Q that is syntactically and semantically correct, relevant to the context and meaningful to answer.

In order to achieve this goal„ we need to train an algorithm to learn the underlying conditional probability distribution

$$P_\theta(Q|X)$$

parametrized by θ. In other words, we can think of this problem as the same one that requires the model learn a function (with a set of parameters) θ during the training stage using content-question pairs so that the probability $P_\theta(Q|P)$ is maximized over the given training dataset.

It is also helpful to frame this problem into a seq2seq learning problem since both the input and the output is most likely a sequence of text character that the model needs to process and learn the relationship from.

4.2.2.1 Case Studies

1. In this paper QG-Net: A Data-Driven Question Generation Model for Educational Content. They use a bi-directional LSTM network to process the input context words sequence. Encoding the answer into context word vectors.

QG-Net generates questions by iteratively sampling question words from the conditional probability distribution $P(Q|C, A, \theta)$ where θ denotes the set of parameters. In order to construct the probability distribution, they first create a **context reader** that process each word c_j in the input context and turns it into a fix-sized representation h_j

Then, they used a **question generator** generates the question text word-by-word, given all context word representation and all question words in previous time steps.

As for the quantitative evaluation, they aimed to minimize the difference between the generated question and the true question in the training set during training. Also, they used the standard back-propagation through time with the mini-batch stochastic gradient descent algorithm to learn the model parameters. They employed teacher forcing procedure for training LSTMs. To enhance performance, they also implemented beam search, a greedy but effective approximation to exhausitively search and select the top 25 cancidate output question sentences. The final one would be the one with the lowest negative log likelihood.

The general QG-Net model Architecture is as below:

2. In this summary Learning to Ask, they used a sentence- and paragraph-level seq2seq model to read text from the input content and to generate a question about the input sentence.

For the second option, we need to encode both sentence and paragraph that sentence belongs to as input, but only attending source sentence hidden states. The performance could be improved with beam search and UNK replacement.

3. In this paper TOPIC-BASED QUESTION GENERATION, they proposed a topic-based question generation algorithm. The algorithm will be able to take in a input sentence, a topic and a question type; then generate a word sequence related to the topic, question type and the input sentence.

They are formulating a conditional likelihood objective function to achieve this goal.

Also, in the paper, they proposed a few frameworks that were used to tackle this problem. The first type is seq2seq model. This model typically uses a bidirectional LSTM as the encoder to encode a sentence and a LSTM as the decoder to generate the target question.

The second approach is question pattern prediction and question topic selection algorithms. It takes in an automatically selected phrase Q and fill this phrase into the pattern that was predicted from pre-mined patterns, which is not done with deep learning.

The last approach is multi-source seq2seq learning which aims to integrate information from multiple sources to boost learning.

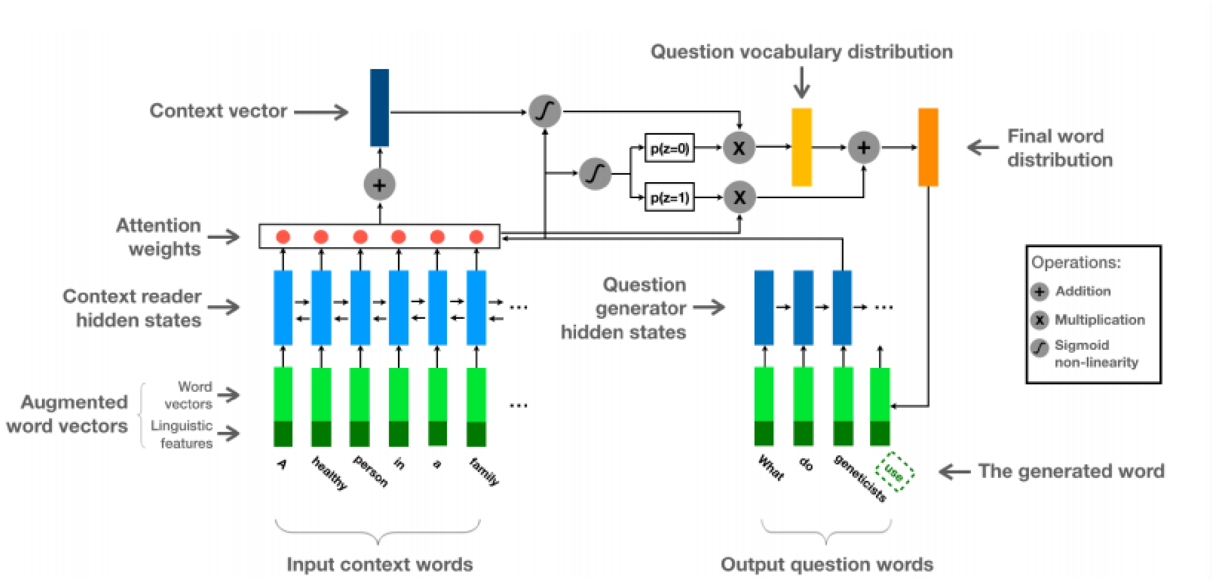


Figure 4.1: ma

4. In this paper A Framework for Automatic Question Generation from Text using Deep Reinforcement Learning they proposed a novel way of solving this problem in which they used a reinforcement learning framework that consists of a generator and an evaluator.

They refer to the generator as the *agent* and the *action* of the agent is to generate the next work in the question. The probability of decoding a word $P_\theta(word)$ gives a stochastic policy.

The evaluator will in turn assign a reward for the output sequence predicted using the current policy of the generator. Based on the reward assigned by the evaluator, the generator updates and improves its current policy. The goal in RL-based question generation is to find a policy that can maximize the sum of the expected return at the end of the sequence generated.

4.2.2.2 Summary

In this QG section, we have discussed 4 algorithms. They provide us a way to frame our problem for which we can apply generative seq2seq model framework. As for our objective function, we are formuating a conditional probability distribution that is conditioned on the provided content (i.e. the video) and answers. Typically, we can use a bi-directional LSTM as the encoder to encode the content and use a LSTM as the decoder to generate the question.

However, as you probabily have noticed that the above examples are focus mainly on processing the text input data instead of videos directly. It demonstrates that more reserch in this new area is needed so as to meet our particular needs.

4.2.3 Answer Generation

4.2.3.1 Close-ended Questions

4.2.3.1.1 Visual Question Answering (VQA)

As what we have covered above, most QG problem focuses solely on generating questions but not the answers based on the context.

VQA is a challenging research problem that focuses on providing a natural language answer given any image and any free-form natural language question. As we are managing to handle the video educational content first that is likely to involve language processing and visual recognition tasks, VQA would be a proper start for us. By leveraging this type of algorithm, we enable our system to easily evaluate the answer provided by learners which could in turn automated the whole question + answering + evaluation cycle.

Since we are dealing with visual input, question-guided attention mechanism is a key component for solving this type of task. Started from the attention mechanism that can adaptively learn the most relevant image regions for a given question. Then to stack multiple question-guided attention mechanisms to learn the attention in an iterative way. Also, it is possible to use bilinear features to integrate the visual features from the image spatial grids with question features to predict attention. Considering the questions in natural language may also contain some noise, the co-attention mechanism can jointly learn the attention for both the image and question.

1. In this paper Deep Attention Neural Tensor Network for Visual Question Answering, they proposed a novel deep attention neural tensor network that can discover the joint correlation over images, questions and answers with tensor-based representation.

As for their workflow, they modeled one of the pairwise interaction (i.e. between image and question) by bilinear features, which is further encoded with the thrid dimension (i.e. answer) to be a triplet using bilinear tensor product. During this step, the model takes in a question + a corresponding image + candidate answers as the input. A CNN (convolutional neural network) a GRU RNN (recurrent neural network) are used for extracting feature vectors and question respectively. Then the representation is passed on as a multi-modal features and integrated by bilinear pooling module. Moreover, they decompose the correlation of triplets by their question and answer types with a slice-wise attention module on tensor to select the most discriminative reasoning process inference.

In the end, they optimize the proposed network by learning a label regression with KL-divergence losses.

They clamined that with these techniques, they can enable scalable training and fast convergence over a large number of answer set.

During the inference stage, they feed the embeddings of all candicate answer into the network and then select the answer which has the biggest triplet relevance socre as the final answer.

The high-level network architecture is as follows:

2. In this paper Question Type Guided Attention in Visual Question Answering, they proposed a model called Question Type-guided Attention (QTA). This model utilizes the information of question type to dynamically balance visual features from both top-down and bottom-up orders.

Finally, they propose a multi-task extension that is trained to predict question types from the lexical inputs during training which generalizes the network into applications that lack question type, with a minimal performance loss.

As for their main contribution, they focus on developing an attention mechanism that can exploit high-level semantic information on the question type to guide the visual encoding process.

Specifically, they introduced a novel VQA architecture that can dynamically gate the contribution of ResNet and Faster R-CNN features based on the question type. In turn, it allows them to integrate the information from multiple visual sources and obtain gains across all question types.

4.2.3.1.2 Video Question Answering

The recent advancements that we discussed above in VQA domain have shown some promising implication. In terms of achieving our particular goal, it is also worth mentioning that VQA might be a good start but it is not sufficient yet. To bridge this gap, let's focus our attention on some video question answering algorithms that have been proposed.

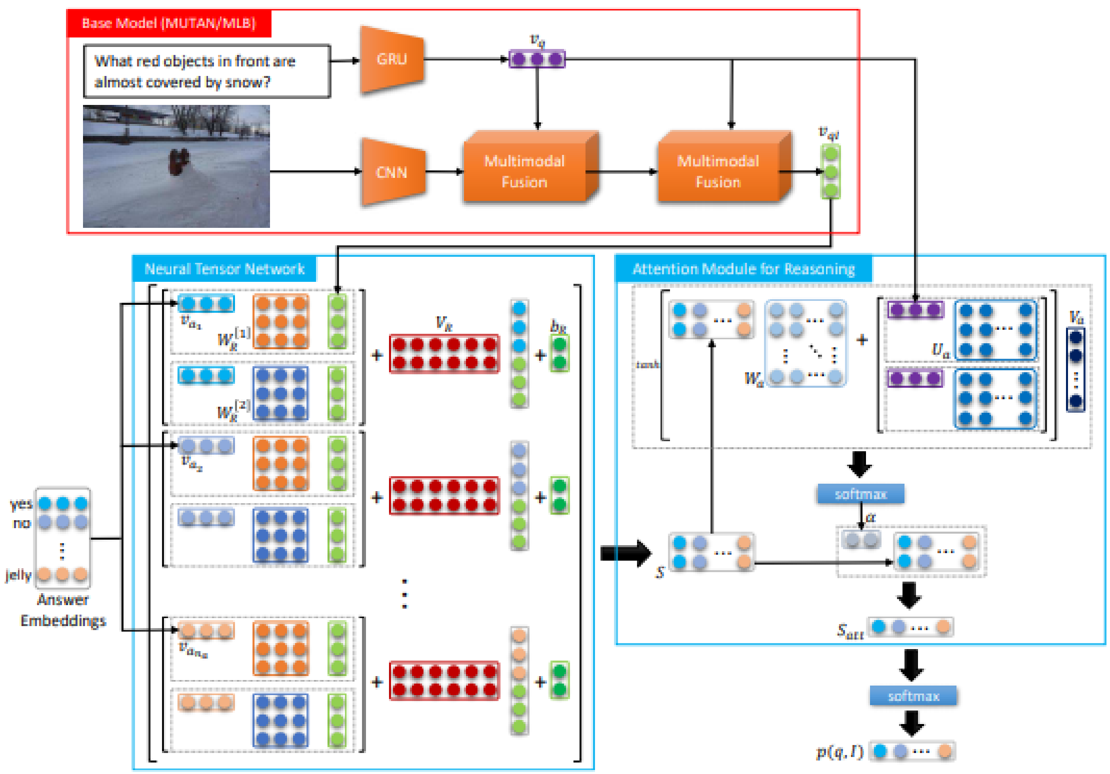


Figure 4.2: Deep Attention Neural Tensor Network

1. In this paper Multi-Turn Video Question Answering via Multi-Stream Hierarchical Attention Context Network, they proposed a hierarchical attention context network for context-aware question understanding by modeling the hierarchically sequential conversation context structure. They also incorporate the multi-step reasoning process fro the multi-stream hierarchical attention context network to enable the progressive joint representation learning of the multi-stream attentional video and context-aware question embedding.

The construct their dataset by collecting the conversational video question answering datasets from YouTube-Clips and TACoS-MultiLevel in which the first one has 1987 videos and the second daaset has 1303 videos. They invite 5 pairs of crowd-sourcing workers to construct 5 different conversational dialogs. In total, they have collected 37228 video question answering pairs for TACoS-MultiLevel data and 66806 ones for YouTube-Clips data.

2. In this paper MovieQA: Understanding Stories in Movies through Question-Answering, they introduced a new dataset called MovieQA dataset that can evaluate automatic story comprehension from both video and text.

They collected 408 subtitled movies and obtained their extended summaries in the form of plot synopses(movie summaries that fans write after watching the movie) from Wikipedia. They used plot synopses as a proxy for the movie. They have annotators create both quizzes and answers pairs by referring to the story plot. Time-stamp is also attached with each question.

In the second step of data collection, they used the multiple-choice answers and question collected as the input to show to the annotators. By doing so, annotators can re-formulate the question and answers while doing the sanity check.

4.2.3.1.3 Summary

By going through the previous examples, we can see that VQA is very particular type of algorithms that is designed to efficiently process image and text input data while making the inference based on the input. Attention is a typical mechanism applied in this type of problems and multiple forms of multipulation on the attention mechanism used in these models have significantly improved the model performance.

Going from VQA to video question answering algorithm, it has been a great leap. The main insight we can directly draw from these video QA papers is that we can follow their steps to collect and annonate our training data by asking crowd-sourcing workers to construct the question and answer pairs. Also, more advanced algorithm like the one described above multi-stream hierarchical attention context network is in need for dealing with video input data in contrast to static pictures.

4.2.3.1.4 Dual Question-Answering Model

Both Question Generaion(QG) and Question Answering(QA) are well-defined 2 sets of models that aim to either infer a question or an answer given the counterpart based on the context. However, they are usually explored separately despite of their intrinsic complementary relationship. In our case, a sysem that can take on both roles simultaneously are needed to fully automated learning + feedback process.

There are some algorithms are designed to fulfill both roles.

1.In this paper Dual Ask-Answer Network for Machine Reading Comprehension they presente a model that can learn question answering and question generation simultaneousely. They tie the network components that playing the similar roles into 2 tasks to transfer cross-task knowledge during training. Then the cross-modal interaction of question, context and answer is captured with a pair of symmetric hierarchical attenion processes.

The high-level architecture of the model is illustrate as below:

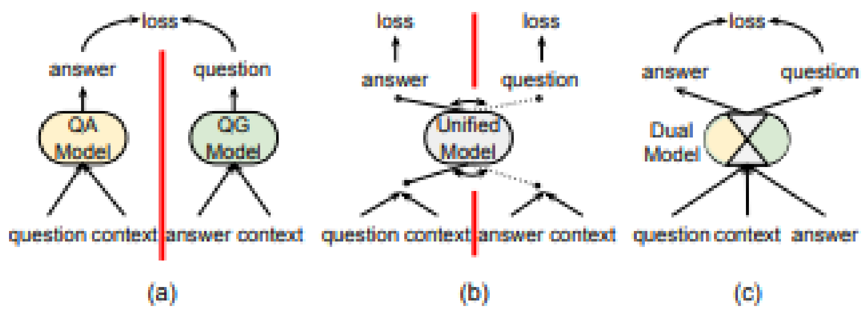


Figure 4.3: Dual Ask-Answer Network 1

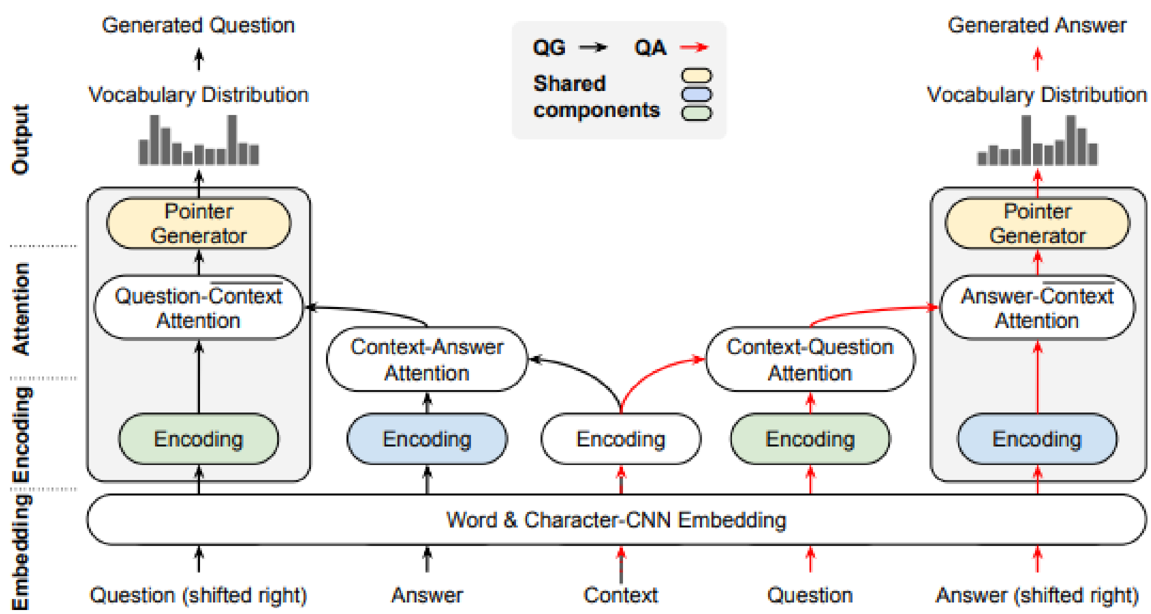


Figure 4.4: Dual Ask-Answer Network 2

In short, the model is composed of embedding layer, encoding layer, attention layer and output layer. The model is fed with a question-context-answer triplet (Q,C,A) and the decoded Q and A from the output layer. Their loss function consists of 2 parts:

- The negative log-likelihood loss
- a coverage loss to penalize repetition of the generated text

2. In this paper Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia, they applied their question-answer pair generation system to 10000 top-ranking Wikipedia articles and create over a million question-answer pairs.

In their task formulation part, they mentioned that they break this task into 2 sub-tasks:

- candidate answer extraction
- answer-specific question generation

To achieve them, they first identify a set of question-worthy candidate answers ans = (A1, A2,...Ai). For each candidate answer Ai, they then aim to generate a question Q - a sequence of tokens y1,y2,...yn - based on the sentence S that contains candidate Ai such that Q asks about an aspect of Ai (of potential interest to a human) and Q might rely on information from sentences that preceeds S in the paragraph. Mathmatically, they compose a function

$$Q = argmax_Q P(Q|S, C)$$

.

3. In this paper Visual Question Generation as Dual Task of Visual Question Answering, they proposed an end-to-end unified model, Invertible Question Answering (iQAN) to introduce question generation as a dual task of question answering to imrpove VQA pefromance.

In achieving their goal, they leverage dual learning framework that is proposed in machine translation area initially, which uses A-to-B and B-to-A translation models to form two closed translation loops and let them teach each other through a reinforcement learning process.

In their VQA component, given a question q, an RNN is used for obtaining the embedded feature **q**, and CNN is used to transform the input image v into a feature map. A MUTAN-based attention module is then used to generate a question-aware visual feature v_q from the image and the question. Later, another MUTAN fusion module is used for obtaining the answer feature $\hat{a}$

4. In this paper A Unified Query-based Generative Model for Question Generation and Question Answering, they propose a query-based generative model for solving both tasks. The model follows the classic encoder-decoder framework. The multi-perspective matching encoder that they are implementing is a bi-directional LSTM RNN model that takes a passage and a query as input and perform query understanding by matching it with the passage from multiple perspectives; The decoder is an attention-based LSTM RNN model with copy and coverage mechanism. In the QG task, a question will be generated from the model given the passge and the target answer, whereas in the QA task, the answer will be generated given the question and the passage. They also leverage a policy-gradient reinforcement learning algorithm to overcome exposure bias (a major problem resulted from sequence learning with cross-entropy loss function).They case both QG and QA tasks into one process by firstly matching the input passage against the query, then generating the output based on the matching results.

As for the training process, they first pretrain the model with cross-entropy loss and then they fine tune the model parameters with policy-gradient reinforcement learning to alleviate the exposure bias problem. During the policy-gradient reinforcement learning algorithm, they end up adopting a similar sampling strategy as the scheduled sampling strategy for generating the sampled output.

4.2.3.1.5 Summary

As what have discussed above, we can see that there have been many attemps taken in the recent years for handling both tasks at the same time and some significant progress have been made.

4.2.3.2 Open-ended Question

4.2.3.2.1 General Question

As I mentioned aboce, one of the most general open-ended questions is to ask learner to provide a short summary of the learning content.

4.2.3.2.2 Specific Question

At the first glance, it appears hard for our system to effectively generate the question. However, this type of task is not that far-fetched by using some specifically designed deep learning frameworks.

As for the automated feedback or learning grading system, there are plenty of suggestions have been proposed as well to tackle such a question. The framework is called automated essay scoring (AES) which focuses on automatically analyzing the quality of writing and assigning a score to the text.

In terms of knowledge or learning evaluation, the format could be diversed i.e. a lecture given by the learner or an short summary essay written by the learner. Regardless the form, we can always convert the content into a predictable text, graphic or audible format that model can process.

As we mentioned above, for these type of task, we can implementing RNN to process the content and even enhance the model performance by adversarially craft input as this paper Neural Automated Essay Scoring and Coherence Modeling for Adversarially Crafted Input illustrated.

4.2.4 Summary of Learning and Feedback Networks

4.2.4.1 summary current research

4.2.4.2 areas where new stuff needs to be made research [current reserach is promising but we need more reserach and innovation in this area]

4.2.4.3 datasets and annotaters needed

Chapter 5

KNOWLEDGE GRAPH

Next, we need to consider how we can select an adequate and relevant learning masterial and generate an effecive learning map for the learners based on their current progress and the general knowledge graph/map, given the ever growing amount of educational content on the web.

As I mentioned earlier, learning is a knowledge accumulation process. Knowledge itself has its unique structure that can help us learn in a most effective and productive way. Knowledge Graph is a great tool that we developed to map and present the structure of knolwedge. In shirt, knowledge graphs are collections of relational facts, where each fact states that a certrain relation holds between 2 entities.

1. In this paper Generalized Embedding Model for Knowledge Graph Mining, they have presented a model for learning neural presentation of generalized knowledge graphs using a novel muli-shot unsupervised neural network model, called the **Graph Embedding Network (GEN)**. This model is able to learn different types of knowlege graphs from a universal perspective and it provides flexibility in learning representations that work on graphs conforming to different domains.

2. In this paper Probabilisic Knowledge Graph Embeddings, they explored a new type of embedding model that can link prediction in relational knowledge graph.

3. In this paper Zero-Shot Question Generation from Knowledge Graphs for Unseen Predicates and Entity Types, they presented a network that can generate question from knowledge b

5.0.1 Summary of Knowledge Graph

5.0.1.1 summary current research

5.0.1.2 areas where new stuff needs to be made research [current reserach is promising but we need more reserach and innovation in this area]

5.0.1.3 datasets and annotaters needed

5.1 KNOWLEDGE JOURNEYS

The knowledge graph is our ground truth and can be applied universally to some extent, but everyone's learning journey is still highly custom. In terms of learning, everyone seems to have their unique set of problems that they are curious about and everyone is on their own mission towards the mastery. As a result, their knowlege journeys could have a lot more degree of freedom depends on the learn's learning history, interests and who they are related to.

We cannot possibly put such an online learning/teaching system into use without taking this crucial factor into our account. However, this is not a trivial problem that can be solved with 1 network or 2.

Let's first formulate our problem before we dive into the possible solutions. Below are few key components that we need to combine to achieve our ultimate goal which is to appropriately guide the learner through their unique knowledge journey:

1. First, we need to

, we can rely on some heuristic and models that have been developed to resolve this type of idiosymcratic issue.

As we all know that a recommender system is an intuitive line of defense against consumer over-choice given the evern growing information available on the web. As we mentioned earlier in the knowledge graph, a authoritative and personalized recommending system is essential for facilitating the learning.

Typically, a recommendation models can be classified into 3 main categories:

1. Collaborative filtering
2. Content based
3. Hybrid recommender system

As I mentioned, here we will mainly focus on hybrid recommender system.

There are a diverse array of achitectual paradigms that are closely related recommending system. Let's take a look at few of them: 1. Autoencoder

2. Convolutional Neural Network
3. Recurrent Neural Network
4. Restricted Boltzmann Machine (RBM)
5. Adversarial Networks
6. Attentional Models (AM)
7. Deep Reinforcement Learning (DRL)

5.1.1 Summary of Current Research and Needs

5.2 Data and Annotation

1. Some datasets used for QG prpblem:

- SQuAD
- RACE

2. Typical datasets used for VQA research:

- VQA dataset that consists of 2 subsets: real images and abstract scenes.
- Compositional Language and Elementary Visual Reasoning (CLEVR) diagnostic dataset that focuses on reasoning.
- Task Driven Image Understanding Challenge dataset (TDIUC); it contains images and annotations from MSCOCO and Visual genome.

5.2.1 Reference

1. A Framework for Automatic Question Generation from Text using Deep Reinforcement Learning
2. Learning to Ask: Neural Question Generation for Reading Comprehension
3. Deep Attention Neural Tensor Network for Visual Question Answering
4. Learning to Ask
5. TOPIC-BASED QUESTION GENERATION
6. Deep Learning based Recommender System

Chapter 6

Conclusion

In conclusion, we presented a new perspective on knowledge acquistion, representation, and proposed an ecosystem that would support a modern and adaptative knowledge ecosystem.

Our main task was to take an individual and begin to get a true depiction of their knowledge beyond their traditional degree which is only a small percentage of someone's education. We focused on taking the world of unstructured educational content online, and how to provide structure in the form of testing and mapping it to a knowledge graph.

There is still much research to be done in bringing to life the EC2QA network as well as the data needed and the collaboration needed amongst machine learning researchers, teachers, and designers needed.

As machine learning researchers we are looking forward to building our next prototype and collaborating around this work.

Chapter 7

About Authors

7.1 Haohan Wang

7.2 Fanli (Christian) Zheng

Chapter 8

Links