TRANSFORMATIONS OF COMPUTATIONAL THINKING PRACTICES IN ELEMENTARY SCHOOL ON THE BASE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES

Gilad Shamir, Ilya Levin

Tel Aviv University (ISRAEL)

Abstract

The importance of acquiring computational thinking skills by a new generation for increasing problem-solving competence is widely recognized. Meanwhile this new generation is growing up in an age of artificial intelligence, which significantly affects how we live, work and solve problems. In order to create a society where citizens have basic computational thinking skills, there should be a powerful elementary school curriculum meeting the requirements of the modern technological society. The main goal of the present study is to bridge the gap between traditional computational thinking skills and new skills required for developing artificial intelligence systems.

Our study deals with elementary school students involved in creating artificial intelligence systems as part of a teacher led course. The course was conducted online due to the Corona virus social distancing of 2020 which resulted in schools closing. The curriculum we developed engages students with a wide range of structured learning activities such as training, testing and evaluating systems, furthermore, modeling features and programming a simple artificial neuron network. For the latter we created a programmable learning environment that enables scaffolded learning activities involving programming. Additional online resources used as part of the curriculum are "IBM Watson" platform and "AI for Oceans" studio by the code.org organization.

As for the research methodology, qualitative and quantitative evaluations were part of the mixed methods research methodology using a variety of evaluation technique, including surveys, semi-structured interviews, in class observations and project evaluations. These were useful for us to refine a framework we suggest for studying and assessing the development of computational thinking for machine learning.

Creating systems which can learn, as a learning activity in elementary school, is at its dawn. Our study is an attempt to contribute to understanding of computational thinking skills associated with constructing artificial intelligence systems. Furthermore, it attempts to suggest a practical implementation of the corresponding curriculum in educational systems.

Keywords: Computational Thinking, Artificial Intelligence, Constructionism, Elementary school.

1 INTRODUCTION

In our days, people use Artificial Intelligence (AI) in their everyday lives and can profit from its recent advances. Nevertheless, one of the most important accomplishments of AI has gone largely unnoticed: the fact that it can supplement human intelligence (Kowalski, 2019). Due to the increased usage of AI machines around us it is only natural that it will find its way into education. Indeed, AI systems for education have grown in numbers in past few years (Domingos, 2015; Klassner & McNally, 2007; Williams et al. 2019).

Children are frequently interacting with artificial intelligent based systems but often have limited knowledge of how they work. This research aims to help unveil this black box using simplified, teachable activities, such that children can learn about the underlying processes of AI and its core computational thinking practices.

A key concept of the proposed research is Computational Thinking (CT). CT can be defined as the thought processes involved in formulating problems and their solutions so the solutions can be carried out by an information-processing agent being a computer or human (Cuny et al. 2010). One way of solving complex problems using CT practices is by using a three-step problem-solving process: 1) decomposing the problem into subproblems, 2) solving the subproblems, and then 3) assembling the solutions of the subproblems into an overall solution (Leiser, 1996). Another three-step problem-solving process is by: 1) abstracting the problem to a simpler one, 2) automating the solution using

ISBN: 978-84-09-17979-4

algorithms so a human or computer can carry it out and 3) analyzing the results in order to verify the original problem was solved (Wing, 2006). These processes require the use of competencies such as Decomposition, Pattern Recognition, Abstraction and Algorithm design. These are now widely accepted as comprising CT and form the basis of curricula that aim to support its learning as well as assess its development (Grover & Pea, 2015). We will refer to these four competencies as `Traditional CT` competencies or `Automation CT` competencies because they are used in a cognitive effort of automating a solution of a problem.

Our study deals with the competencies required for constructing an artificially intelligent system and whether they are different from the traditional CT competencies. An AI system is a computer system that behaves in a way that would be called intelligent if a human were so behaving (McCarthy et al. 2006). AI is a wide-ranging topic and comprises several techno-scientific branches (Corea, 2018) one of them being machine learning which will be discussed later in this section.

In (Domingos, 2015) it is claimed that AI in general, and machine learning in particular, is the scientific method on steroids. It will change every field, from health care to politics to journalism. In other words, the process we have called “automation” will be automated itself. “The industrial revolution automated manual work and the information revolution did the same for mental work, but machine learning automated automation itself” (Domingos, 2015). In the final clause of this sentence, Domingos imagines society running on the steam of some other intelligence.

Let us analyze this statement of AI being a form of "Automation of automation" from a computer software point of view. Working in the traditional `Automation` paradigm, the software architect intends for the system of agents (a.k.a. objects) to be inner controlled and self-regulated, meaning the programmers set rules for the agents and their inter-relations. During runtime, e.g. when the software executes, these agents act by themselves with no human intervention. However, this is not the case in the "Automation of automation" paradigm, where the system of agents autonomously changes the behavior of the system in a way that their creator (e.g. the programmer) is not aware of and cannot surely predict. In such cases not only that once the program is executed there is no human intervention, also the behavior of the agents changes with no intervention.

One of the interesting aspects of constructing AI systems is that it challenges the main CT practices. AI does not follow the three classical CT problem-solving steps mentioned above. In this sense, AI system construction makes problem solving different and could perhaps redefine educational goals with regards to CT as a set of problem-solving competencies. With this in mind, we go back to the claim that AI is "automation of automation" and define the following question which is the essence of this research: What are the computational thinking skills a person should have in order to create an automation of automation solution to a problem?

With regards to education AI systems have exploded too suddenly for education to keep up, and it has a reputation for being a difficult subject to learn (Domingos, 2015). It is worth noting here that the potent idea of AI in education pre-dates the 21st century. An early attempt to teach children about AI came about in 1971 in which researchers wanted children to explore AI through LOGO programming (Papert & Solomon, 1971). Today, researchers are just beginning to implement some of those ideas (Williams et al. 2019). Although the idea of AI in education is popular and is 50 years old, much of the research was focused on ethics of AI or use of AI as a tool to personalize student progression in learning. For instance, the Intelligent Tutoring Systems paradigm has dominated the field of AI in education since the 1970's because it formed the answer to the supposedly unintelligent computer-based learning systems which had prefixed teaching scenarios (McArthur et al. 2005; Andriessen & Sandberg, 1999; Lucking et al. 2016). Currently, there is still scarcity of research on students involved in constructing AI systems and lack of a formal framework for studying and assessing the development of computational thinking for constructionist AI learning.

We chose to use the constructionist pedagogy in our curriculum. Constructionist pedagogy, sometime referred to as learning-by-design, applies to all domains and suggests that students learn best when they engage in the creation and usage of external representations for modeling and reasoning (Blikstein & Wilensky, 2009; Papert & Harel, 1991). A great deal of research has been done on designing introductory programming environments that make the task of programming more accessible to younger learners as well as more fun and engaging (Kelleher & Pausch, 2005). One approach that has gained popularity in recent years is visual programming environments that provides visual blocks as commands which can be connected. This approach has been found to be an effective way of allowing novice learners to have early programming successes (Maloney et al. 2008). However, a general-purpose programming environment has its flaws. It has endless opportunities and is a time-consuming process for a learner to reach the required domain, thus making it impractical for a wide use (Guzdial, 2010).

To overcome this flaw, researchers create a domain specific microworld, based on a general-purpose programming environment having a dedicated set of blocks for the students to use (Jona et al. 2014; Sabharwal et al. 2020). For enabling the course participants to program a simple artificial neuron network, a basis of some AI systems, we created a microworld which is a programmable learning environment based on MIT Scratch (Resnick et al. 2009) that enables scaffolded learning activities. Another microworld which is part of the curriculum. is the IBM Watson platform (Sabharwal et al. 2020).

Some of the microworlds used in our curriculum are based on machine learning. Machine learning (ML) is a field of artificial intelligence that uses statistical techniques to give computer systems the ability to learn to categorize data without being explicitly programmed. In machine learning, a person programs a computer software to use example data to solve a given problem (Alpaydin, 2020). Once programmed, the system requires to be trained based on given data. The trained ML model can later be used to predict the probability of a future event within an acceptable reliability. For example, predicting if any given image is that of a cat requires using images of cats and images without cats to train a model using machine learning techniques such as tagging each image with a label of "cat" and "no cat". Later by feeding a new image to the model it predicts its probability being a cat.

With supervised learning, a machine learning process is a two-step process which requires some intermediary steps (Guyon et al. 2006; Khalid et al. 2014). In the 1st step called "Training" one must apply appropriate data preprocessing such as data acquisition, feature selection and data splitting. These make the dataset appropriate for machine learning. One must then perform algorithm selection and activate it on the preprocessed data set. In the 2nd step called "Validating" (a.k.a "Predicting" or "Scoring") one must test the model to maximize the predictive performance of the final system. This requires gathering a validation data set which was not part of the training data set, predicting the result for each data entity running the system and validating the system`s predictions. See Figure 1.

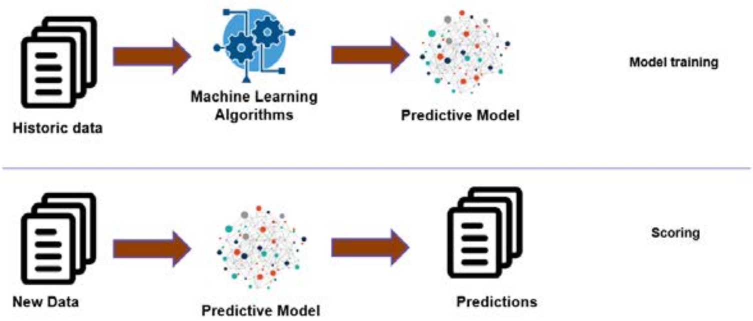

Figure 1. Machine learning - two-step process

2 RESEARCH QUESTIONS

In a previous section Computational Thinking (CT) competencies of Decomposition, Pattern Recognition, Abstraction and Algorithm design were referred to as "Automation CT" or "Traditional CT" competencies because they assist in automating a solution. In the age of AI, claimed to be "Automation of automation", and with the constructionism pedagogy gaining popularity, it is important to identify the competencies of problem solving relevant to constructionist AI activities. Meaning, to identify the computational thinking skills required for constructing an AI system.

Our study deals with elementary school student's computational thinking skills as they are expressed in unique constructionist AI activities. Accordingly, the research questions are:

1. How effective is the constructionist approach for introducing in elementary school the construction of artificial intelligent systems?
2. Should the traditional computational thinking set of competencies be enhanced in the realm of student led construction of artificial intelligent systems in education? If so, how?

3 METHODOLOGY

This research studies student's computational thinking (CT) competencies as they undergo a novel educational program: a constructionist Artificial Intelligence (AI) course for elementary school children. With regards to this goal, the research design is the exploratory sequential design (Creswell & Plano Clark, 2011). It includes two parts: (1) collection and analysis of qualitative data; (2) building from the exploratory results quantitative part to test and generalize the initial findings, and then an interpretation of how the quantitative results build on the initial qualitative results. Different methods were used to assess the different study components and to assess the plausibility of identified threats to validity. A detailed description of each instrument is displayed later in this section.

Prior to creating the course assessments, we investigated AI and machine learning (ML) literature and created a list of relevant computational thinking competencies suitable for young children. They are based on the skills required to train and validate a ML system and relevant concepts. These were reevaluated during the course and adjusted according to our findings.

To create a taxonomy for CT for constructionist ML learning we embedded our findings in the "Computational thinking assessment framework" (CTAF) (Brennan & Resnick, 2012) thus forming a framework for studying and assessing the development of computational thinking for constructing machine learning systems (CT-ML). We used the key dimensions of CTAF: 1) Computational concepts which are the concepts practitioners engage with as they create computational artifacts, 2) Computational practices which are the practices practitioners develop as they engage with the concepts, 3) Computational perspectives which are the perspectives designers form about the world around them and about themselves. CT-ML forms the taxonomy for this research and is further described in the results section of this paper.

The pilot study was conducted with seven 6th grade students from the "Hashfela" region in Israel. 6 boys and 1 girl volunteered to participate in the 12-hour, teacher guided, online AI course we conducted. Every participant and their parents gave informed consent to participate in the study. The original plan was to conduct the study in a school during school hours, but the Corona virus social distancing of 2020 resulted in schools closing thus moving to online learning. The 7 students were divided to 2 separate groups. A group of children in a course consists of up to 4 students accompanied with a facilitator.

3.1 Assessment

At the beginning of the course a validation was carried out of participants' basic acquaintance with 3 fundamentals: 1) Online conference tool acquaintance: Activating the microphone and camera, using multiple tabs in an internet browser and chatting 2) Programming with MIT scratch acquaintance: Students with insufficient background are given an introductory programming assignment. 3) Artificial Intelligence (AI) system acquaintance: A presentation of what is AI and example of systems that implement AI it like "google translate" and "YouTube" which are commonly used by children is Israel.

After we validated the students have the basic knowledge a pre course assessment was done. This was followed by the main part of the course - a set of learning modules. Each module was comprised of a set of activities alongside a variety of assessments. A typical lesson duration was 2 hours with 6 lessons carried out. Following the course, a post course questionnaire was administered. Finally, semi structured interviews were held with 1 or 2 participants only. These were transcribed for further analysis.

During the online course, the children performed the assessments individually, from their home, using their own computers. A variety of online tools were used to capture children's thoughts such as videotaping the lessons and interviews. An online shared bulletin board tool was used to post ideas and discuss them. This also assisted in gathering research data. The online chat system was also used.

Henceforth is a detailed description of each instrument:

Instrument 1 - Course level questionnaire: Pre and post questionnaires were administered. They included closed and open-ended questions about children's perspective of AI, knowledge about ML process, CT competencies and desire to learn about AI. The prequestionnaire consisted of 19 questions while the post was comprised of 32 questions. The additional questions investigated enjoyment of course activities, self-efficacy with AI construction and in-depth computational thinking AI competencies.

The questions related to CT competencies were coded so each question targeted a specific competency. Due to the novelty of the topic it was necessary to thoroughly prepare the questions and continuously refine them as the course progressed and insights on CT-ML increased. Henceforth is the questionnaire creation process:

1 Create a set of questions targeting CT-ML taxonomy.
2 Different reviewers, who understand well CT-ML, mapped the questions to the taxonomy.
3 We examined student's ability to understand the questions and refined them as needed.
4 Add and refine questions as CT-ML was enhanced.

An example of one of the questions: A machine learning system trained to recognize pictures showing a crosswalk should be good at which of the following tasks: a) Identifying pictures of crosswalks b) Identify pictures of a person crossing a road c) Identify pictures of fruits d) Identify pictures of traffic lights.

Instrument 2 – Feature selection questionnaire: Data analysis is a major challenge for practitioners in the field of machine learning (Cai et al., 2018). Feature selection provides an effective way to aid in this challenge by improving learning accuracy and facilitating a better understanding of the data set. Feature selection is the task of choosing a small subset of features for building the prediction model.

Pre and post questionnaires were administered before learning module 1 and after it. The feature selection questionnaires asked students to identify features in a data set of fish images (see Figure 2). It does so by asking to list similarities and differences between the different images. For example: All images are similar in that they have 2 eyes, but they are different because the eyes do not point to the same direction.

Figure 2. Example of a data set of fish images

Instrument 3 - Feature map modelling: The feature map task requested each child to decompose the task of feature selection into 2 parts. 1st of all create a mind map of body parts. Only than iteratively go over each body part and add to it its features. A valid feature was explained to be a dichotomic attribute of the images. For this task the students used an online, easy to use tool, www.mindmup.com. The maps were later used for analysis.

Instrument 4 – Semi structured interviews:

The interviews were conducted in small groups of up to 2 children at a time. This was done in order to give ample opportunity for each of the participants to express their thoughts in detail. Interviews were 30 minutes long. We prepared a list of predetermined questions but allowed the interviews to unfold in a conversational manner offering participants the chance to explore issues they feel are important.

3.2 Course modules

The AI curriculum consists of four modules: 1) "What is AI", 2) "Teaching an image-based AI system", 3) "Constructing an artificial neuron network", 4) "Teaching and programming a text-based AI system". This paper describes modules 1 and 2. In future studies modules 3 and 4 will be described. The course curricula and assessments covered a set of steps common to ML construction : 1) Category selection, 2) Data selection, 3) Feature selection, 4) Data filtering, 5) Split to data sets, 6) Training, 7) Validation. A traditional part of the ML process, "Algorithm selection", was omitted due to high math requirements.

Description of module 1 - What is AI

During this module participants are involved in discussion how AI works and how is it different than algorithm-based programming. The purpose is to expose them to what we referred to in an earlier section: the difference between "Automation" and "Automation of Automation" paradigms. For instance, they are asked to each open the YouTube website and explain why it recommends different videos for each other. They also are given 3 tasks of programming a figure to reach a goal.

Description of module 2 - Teaching an image-based AI system

In this module the children perform a set of activities of training an AI system to classify images into two categories. The categories change between the activities with increased complexity. These activities require the student to train, test and validate the classifier model. The context of the module is "Ecological systems" based on the 6th grade science curriculum in Israel ("Ministry of Education science curriculum", n.d.).

In activity 1 the children train, test and validate an AI program to recognize if an image is a fish or not. They are faced with 2 kinds of pictures: fish or garbage. In activity 2 they are faced with 3 kinds of pictures: 1) non-fish sea creatures, 2) fish and 3) garbage. Their task is to train the AI system to recognize if the picture shows an agent that belongs in the ocean. In activity 3 the children select a category of their own from a list of over 20 categories. The categories vary in classification difficulty from "is happy" and "is hungry" to "is silly" and "is glitchy".

This module uses "AI for oceans" studio (studio.code.org/s/oceans) microworld.

4 RESULTS

In this paper we discuss a taxonomy for teaching ML which is based on our research as well as student's satisfaction from the course and their interest to learn more. Finally, a detailed analysis of the "feature selection" computational skill is presented. Additional results will be part of future papers.

4.1 Taxonomy

One of the research goals was to invesigate if traditional computational thinking practices should be enhanced with use of artificial intelligence, and if so, How?

Through the course and interactions with the children, as they practice constructionist AI, we were able to refine our initial taxonomy, thus forming a "Machine learning computational thinking studying and assessing framework for constructing machine learning systems" (CT-ML). The framework is shown in Figure 3.

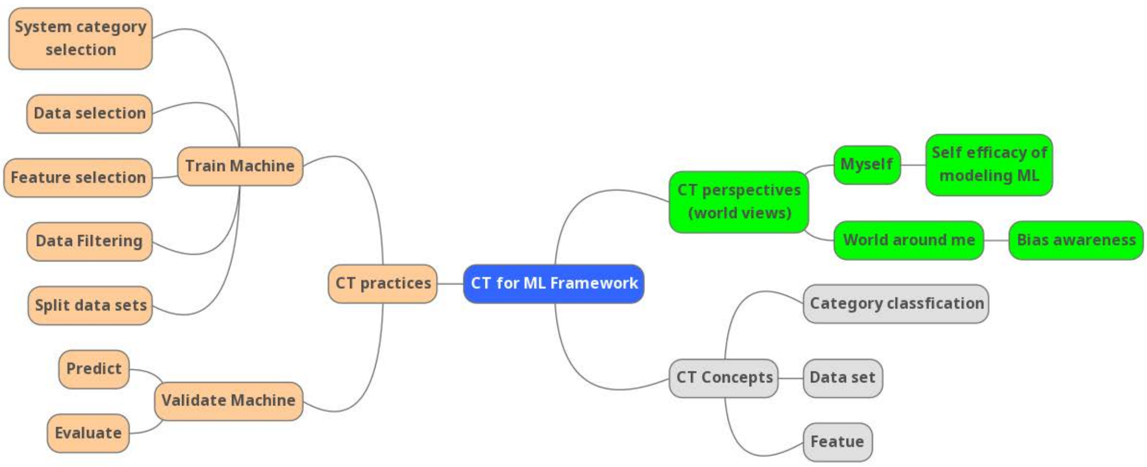

Figure 3. Framework for studying and assessing the development of computational thinking for constructing machine learning systems (CT-ML)

In this figure, the orange coloured items are the computational practices. The practices have been split into the skills required for each step of the ML constructing process: 1) Training the system and 2) Validating its classification model. The green colored items are the computational perspectives dealing with how one perceives himself and his surrounding in the light of constructing ML systems. In grey are the computational concepts which are what practitioners engage with as they create ML systems.

4.2 Motivation

We used the ARCS Model of instructional design (Keller, 1987) which is based on empirical investigations, as a tool for measuring the student's motivation. The model contains four conceptual categories that “subsume many of the specific concepts and variables that characterize human motivation”. Categories are: 1) Attention – should be sustained during the learning process, 2) Relevance – relates to why is this material important to me, 3) Confidence - can influence a student’s persistence and accomplishment, 4) Satisfaction makes people feel good about their accomplishments.

All children asked at some point or another during the course "Will be a contingency course they can join", because they wanted to learn more. The questionnaires showed that as well. The participants highly valued the courses contribution to their understanding of AI averaging 4.6 out of 5 on a liker scale (n=7, Std. dev.=0.7). They thought AI was highly interesting (n=7, avg=4.4 of 5, Std. dev.=0.8). A student that gave the course the lowest rating, 3, was asked about the discrepancy between the verbal enthusiasm and the written one. He explained the 3 as incidental: "I gave a low rate on the satisfaction questionnaire because I don't know why WebEx online meeting was used when I know of a much better one. This is why the course can be improved".

All 7 students participated in all the meetings, not missing even one, they scored high on the question about their confidence to deal with AI. When asked what they thought about the course, some remarks was related to satisfaction of basic pedagogical principles – continuously ask them what they think and use a variety of ways for them to express their thoughts such as the online shared bulletin board and the online chat system. They expressed that they have discussed the course with friends or siblings, which can also be an indication they feel its relevant to them.

These give a strong indication of the student's high motivation to learn about AI as part of the course.

4.3 Feature selection practice

According to ML literature and according to our findings, one of the most challenging practices for practitioners is "Feature Selection". Feature Selection (a.k.a feature extraction) is a process of finding only those attributes that contain the relevant information for ML system`s data classification.

Results of feature selection skills were gathered using various assessment instruments including questionnaires, tasks and interviews and in class observations. We created several measurements for analysing feature selection: 1) Number of features, 2) Description level, 3) Validity ratio, 4) Richness ratio, 5) Error types, 6) Self-efficacy.

One type of questions was related to a data set of fish images. For instance, a question asked them to write the identical attributes of all fish and another to write the attributes which were different. For example, they all had a pair of eyes, so these are common among all fish, but the eyes had 2 different looks, so this differentiates them. We'll now look into 2 measurements using 2 examples: 1) A participant wrote in the pre-questionnaire: "They all have the same background" which is 1 attribute described by 6 words. 2) A participant wrote in the post-questionnaire: "They all have a mouth, they all have one central body and the other parts come out from it, they all have fins and everyone has a mouth pointing to the right". This description consists of 32 words and 3 attributes (mouth direction, body and fins).

Our analysis of all the participants answers showed that the description level between the pre and post-tests went up from an average of 13.1 words per participant to 17.1. On the other hand, the number of attributes between the pre and post-tests dropped from an average of 7.7 attributes per participant to 6.6. We don't have an additional supporting measurement for this drop. All other assessment results indicated an incline in feature selection capability. Since that every attribute was described in more detail in the post questionnaires than, perhaps, the drop was due to reasons unrelated to the feature selection skill but rather to other factors like that students found it hard to type so much. Meaning the more detailed their answer was for a certain attribute the less inclined they were to add more attributes to their list. Furthermore, both results of average attributes count (pre-7.7, post-6.6) were low because the images had many more attributes.

Another assessment method to measure feature selection competency was the feature map modeling assessment. In this task participants were asked to create a feature map by first dividing the fish figure to body parts, adding each of them to the map. Only after that part was over were asked to look at one body part at a time and for it to add the selected features to the map. A part of a feature map created by one of the students is shown on Figure 4.

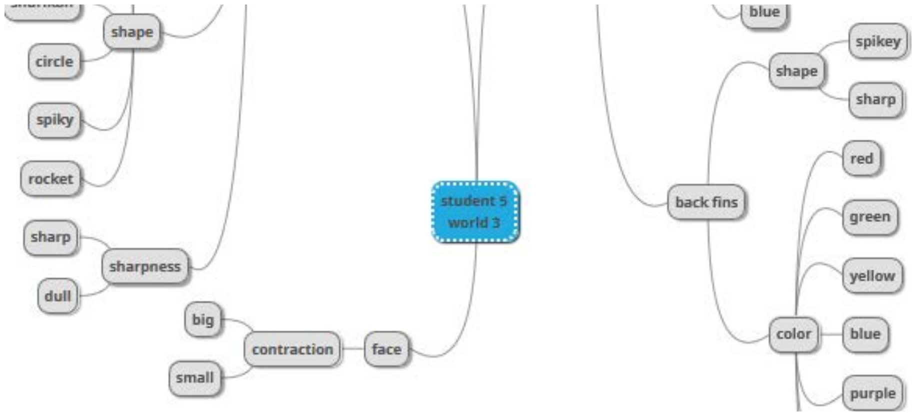

Figure 4. Feature map

In this task 7 participants attended but one map failed to save so we analysed 6 maps (n=6). The average number of features in this task was 27.33 (Std. dev.=11.87) which is considerably higher than the 6.6 average on the written test. We also noticed the standard deviation was high showing that the results were spread out. We further analysed the maps to find the validity ratio, richness ratio and each feature the students added to their map was investigated in order to decide if it is valid.

The validity ratio was calculated as the valid features divide by the total features the participant had on their map. Validity rate was very high (avg = 94%, Std. dev. = 0.13) meaning most features added to the map were indeed valid features and have met the criteria.

In order to find the richness ratio, we looked at the student with the most valid features as a 100% richness (42 features) and calculated the others rate with relevance to that student. The richness rate has averaged 62% with a standard deviation of 30% indicating that the diversity was large with 3 students below average and 3 students above average.

Next, we looked at participants self-efficacy with their ability to train a AI system. In the feature modeling task the participants were given 3 data sets of fish images with increased variance of details. For example, data set #1 had 3 mouth types while data set #3 had 7 mouth types. 5 participants chose the most difficult level while 2 more chose the 2nd most difficult one. This indicates a high level of self-efficacy of feature selection. This assumption was strengthened by the post course questionnaire analysis in which on a 5 level Likert scale they had an average of 4.6 (n=7, Std. dev.=0.7) on the statement "I believe I can train an AI system" and an average of 4.5 (n=7, Std. dev.=0.5) on the statement "It is clear to me how an AI system learns".

4.4 Understanding AI processes

Using a questionnaire and interviews, we collected data on what participants consider as key to good ML system training and evaluating. A relevant questionnaire was administered during the middle of the course, after they started to understand the AI lingua. It was administered again at the end of the course. For example, they were given 5 statements to mark those they believe are fundamental to the training process. The average correct answers inclined from 3 out of 5 (n=7, Std. dev.=1.3) to 4.4 (n=7, Std. dev.=0.9) When asked to explain their considerations some of the explanations showed a good understanding of AI training. Two examples: "Feature selection is important because if you have good features it makes the fish image be unique" and "If you have more than one person to tag the image you get more than one opinion and it is less likely to have the AI system make a mistake".

The participants were able to explain the following terms: Neural network, ML training and Machine learning. They demonstrated understanding possible implications of AI and suggested how to decrease the risk. For instance, a participant said " autonomous cars need much training… maybe even 1000 hours. They need to train again and again until it learns". Another participant added: "If you want the car to see the traffic signs you need to place the signs in many places even in the ocean, not just one place, so it will understand".

They showed good understanding of invalid predictions and how they relate to the training process. For example when asked why did a face recognition software identify a girl with dotted pajamas as having 3 faces they said: "The prediction was wrong because the pajamas had dots which looked like face so it "saw" something that looks like a face and when it looked they found it. It thought 2 dots were the eyes and because the lower dots together looked like a smile it thought it was the mouth".

5 CONCLUSIONS AND FUTURE WORK

This paper presents findings of a pilot study on computational thinking competencies of young students who undergo a constructionist AI course. Through constructing, training, testing and validating AI systems, children learn how a human can create a computerized system to become "intelligent" using input data. They understand how data can be manipulated for the system to learn differently therefore after retraining it the predictions change. Image based data, along with a constructionist pedagogy, was found to be effective in teaching children about ML concepts and developing students' CT skills. The students had demonstrated high engagement with course materials and high motivation to learn. They also demonstrated an enhancement in the computational thinking ability that was investigated.

The studies limitations are that it was conducted with a small group of students, which all chose to participate. While such small-scale qualitative studies are useful for in-depth exploration of a phenomena, they do not allow for generalization beyond the sample under investigation (Creswell, 2009). Course content dealt with image-based AI only so perhaps text-based or voice-based AI may have different results. The course was online and not in a regular class mode.

This study includes key design considerations for future development of the proposed approach of AI learning. In future studies, we plan to expand the constructionist AI learning with regards to computational concepts, practices and perspectives according to the CT-ML framework we proposed. With regards to computational practices we plan to create assessments for category selection, data set split and data filtering skills. The relation between traditional CT and AI based CT is a key factor for the future of CT in education; therefore, we plan to further analyze it. Regarding the neuron based programmable learning environment, we plan to study its applicability in two dimensions: Ease of use and Project richness. Finally, we would like to add an additional analysis instrument for student posed questions types and their relation to thinking skills. We believe that our study opens up promising prospects for the constructionist AI learning approach in elementary school.

REFERENCES

[1] Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

[2] Andriessen, J., & Sandberg, J. (1999). Where is education heading and how about AI. *International Journal of Artificial Intelligence in Education*, *10*(2), 130-150.

[3] Blikstein, P., & Wilensky, U. (2009). An atom is known by the company it keeps: A constructionist learning environment for materials science using agent-based modeling. *International Journal of Computers for Mathematical Learning*, *14*(2), 81-119.

[4] Brennan, K., & Resnick, M. (2012, April). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada* (Vol. 1, p. 25).

[5] Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70-79.

[6] Corea, F. (2019). AI Knowledge Map: how to classify AI technologies. In *An Introduction to Data* (pp. 25-29). Springer, Cham.

[7] Creswell, J. W. (2009). Mapping the field of mixed methods research.

[8] Creswell, J. W., & Plano Clark, V. L. (2011). Designing and conducting mixed methods research. *2 ed.*

[9] CSTA & ISTE (2011). Operational Definition of Computational Thinking for K–12 Education. Retrieved from http://csta.acm.org/Curriculum/sub/CurrFiles/CompThinkingFlyer.pdf

[10] Cuny, J., Snyder, L., & Wing, J. M. (2010). Demystifying computational thinking for non-computer scientists. Unpublished manuscript in progress, referenced in http://www. cs. cmu. edu/~ CompThink/resources/TheLinkWing. pdf.

[11] Domingos, P. (2015). The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books.

[12] Grover, S., & Pea, R. (2013). Computational thinking in K–12. A review of the state of the field. Educational Researcher, 42(1), 38-43.

[13] Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. In *Feature extraction* (pp. 1-25). Springer, Berlin, Heidelberg.

[14] Guzdial, M. (2010). Does contextualized computing education help? *ACM Inroads*, *1*(4), 4-6.

[15] Jona, K., Wilensky, U., Trouille, L., Horn, M. S., Orton, K., Weintrop, D., & Beheshti, E. (2014). Embedding computational thinking in science, technology, engineering, and math (CT-STEM). In *future directions in computer science education summit meeting, Orlando, FL*.

[16] Leiser, D. (1996). Constructivism, epistemology and information processing. *Anuario de psicología/The UB Journal of psychology*, (69), 93-114.

[17] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education.

[18] Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys (CSUR)*, *37*(2), 83-137.

[19] Keller, J. M. (1987). The systematic process of motivational design. *Performance+ Instruction*, *26*(9-10), 1-8.

[20] Khalid, S., Khalil, T., & Nasreen, S. (2014, August). A survey of feature selection and feature extraction techniques in machine learning. In *2014 Science and Information Conference* (pp. 372-378). IEEE.

[21] Klassner, F., & McNally, M. (2007). Demonstrating the Capabilities of MindStorms HXT for the AI curriculum. In *American Association for Artificial Intelligence*.

[22] Kowalski, R. (2011). Computational logic and human thinking: how to be artificially intelligent. Cambridge University Press.

[23] Maloney, J. H., Peppler, K., Kafai, Y., Resnick, M., & Rusk, N. (2008, March). Programming by choice: urban youth learning programming with scratch. In *Proceedings of the 39th SIGCSE technical symposium on Computer science education* (pp. 367-371).

[24] McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Magazine, 27(4), 12.

[25] Ministry of Education Science curriculum. (n.d). Retrieved from http://cms.education.gov.il/ EducationCMS/Units/Mazkirut_Pedagogit/MadaTechnologya/tochnitLimudim/yesodi+tl.htm

[26] Sabharwal, N., Barua, S., Anand, N., Aggarwal, P., (2020). Developing Cognitive Bots Using the IBM Watson Engine. APress.

[27] Papert, S., and Solomon, C. 1971. Twenty things to do with a computer. Cambridge, MA.

[28] Papert, S., & Harel, I. (1991). Situating constructionism. *Constructionism*, *36*(2), 1-11.

[29] Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., ... & Kafai, Y. (2009). Scratch: programming for all. *Communications of the ACM*, *52*(11), 60-67.

[30] Williams, R., Park, H. W., Oh, L., & Breazeal, C. (2019, July). Popbots: Designing an artificial intelligence curriculum for early childhood education. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 9729-9736).

[31] Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35.