Toward a User-Adapted Question/Answering Educational Approach

Chiara Alzetta
DIBRIS, University of Genoa, Italy
chiara.alzetta@edu.unige.it

Giovanni Adorni
DIBRIS, University of Genoa, Italy
adorni@unige.it

Ilknur Celik
Cyprus International University
Cyprus
ilknur.chelik@gmail.com

Frosina Koceva
DIBRIS, University of Genoa, Italy
frosina.koceva@edu.unige.it

Ilaria Torre
DIBRIS, University of Genoa, Italy
ilaria.torre@unige.it

ABSTRACT

This paper addresses the design of a model for Question/Answering in an interactive and mobile learning environment. The learner's question can be made through vocal interaction or typed text and the answer is the generation of a personalized learning path. This takes into account the focus and type of the question and some personal features of the learner extracted both from the question and prosodic features, in case of vocal questions. The response is a learning path that preserves the precedence of the prerequisite relations and contains all the relevant concepts for answering the user's question. The main contribution of the paper is to investigate the possibility to exploit educational concept maps in a Q/A interactive learning system.

CCS CONCEPTS

• **Applied computing** → **Interactive learning environments**;
• **Information systems** → *Personalization*;

KEYWORDS

Education, Personilized Learning Path, Q/A, MOOC, educational concept map

ACM Reference Format:
Chiara Alzetta, Giovanni Adorni, Ilknur Celik, Frosina Koceva, and Ilaria Torre. 2018. Toward a User-Adapted Question/Answering Educational Approach. In *UMAP'18 Adjunct: 26th Conference on User Modeling, Adaptation and Personalization Adjunct, July 8–11, 2018, Singapore, Singapore.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3213586.3226214

1 INTRODUCTION

Nowadays learning platforms like Massive Open Online Courses (MOOCs) are used widely as tools for supporting distant learning. The underlying philosophy of MOOCs is based on providing learning content accessible from home, but the increasing use of mobile devices and web applications made learning even more ubiquitous, allowing MOOC users to access educational content anytime and everywhere. Although many still use learning platforms from their home, we observe an increasing use of tablets, smart phones and smart televisions [20]. This trend allows us to consider new possibilities when designing courses and activities for distant learning; vocal interaction is one of those. In principle every mobile device offers the possibility to interact via voice in addition to using manually typed text. MOOCs can take advantage of this to widen the number of contexts of use: for instance, to access a course from smart televisions, vocal interaction is almost necessary.

A well-known drawback of MOOCs is that they are designed to be "massive". Many studies focus on strategies to make MOOCs more personalized addressing different needs of the learners but remaining pedagogically efficient [25]. A MOOC course generally tackles many concepts, trying to offer a view that is complete as much as possible on a subject. Sometimes it may happen that the user is not interested in the entire course but only on sections addressing a specific issue, so viewing all the contents of the course is not an efficient way to obtain her/his goal. This causes the users to abandon the course and look for information elsewhere, where they can select what corresponds to their needs. If a user was able to express a specific need to the platform, then it would be possible to create a personalized learning path containing only the learning materials related to her/his interests.

We address this problem using Question/Answering (QA). QA is the task of automatically providing appropriate answers to questions asked in natural language, searching it in documents or a form structured knowledge [18].

In our view, the QA task should be considered as learning path retrieval: the platform returns a personalized sequence of contents pertinent to the information the user sought, but differently from classical information retrieval, the sequence is organized in such a way that it also has an educational value. In an educational environment one concept is not disjoint from the others: each concept has dependency relations with others, like prerequisite. Therefore the content of a course can be organized in an Educational Concept Map (ECM), a formal representation of the domain of knowledge that explicits those relations. A structure like this allows to retrieve the most appropriate learning path without leaving out prerequisite concepts [1].

In the field of education, QA Systems (QAS) are commonly used as virtual tutors, providing assistance when learners struggle to find answers or in addressing common misconceptions [2]. The questions a learner asks to the QAS not only identify the concepts

ACM ISBN 978-1-4503-5784-5/18/07.
https://doi.org/10.1145/3213586.3226214

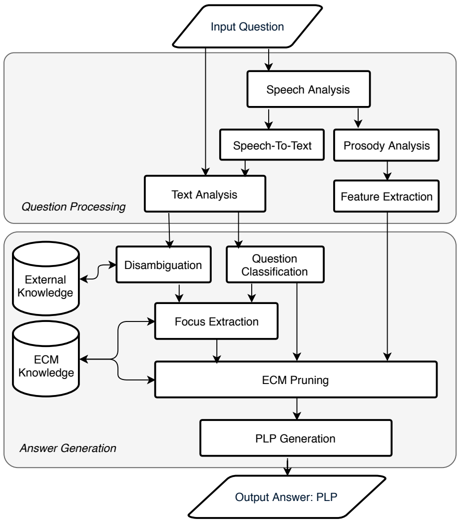

Figure 1: Q/A Flow Model

the learner needs to understand, but, if vocal interaction is allowed, they may contain useful clues about the confidence of the learner on that topic, together with her/his emotional state. The literature shows that acoustic prosodic elements may provide information about several features of the user, including emotions and engagement [15, 21]. Based on this literature, our model includes the use of acoustic prosodic features to select from the ECM, and connect, the learning concepts and materials that most satisfy the user query.

The aim of this paper is thus to address the design of a model for personalized learning path generation accessible through vocal interaction or typed text. The learning path is based on the question focus and type and it is also personalized by considering the confidence of the learner with the topic and her/his engagement extracted both from the question and prosodic features of the speech.

The paper is organized as follows: first the model is described in section 2, then the types of questions we can receive from users are presented in section 3. We describe the process of building the personalized learning path in section 4 and discussion and related work can be found in section 5.

2 Q/A FLOW MODEL

The QA model is designed to have a natural language user interface so the learner can ask questions in natural language and obtain a learning path containing the information pertaining to the question as answer. The flow model, described in Figure 1, starts from the input question and returns the Personalized Learning Path (PLP).

This model has two main components: a question processing block, which includes tasks to analyze the input in order to extract useful information to be used in the subsequent block, and an answer generation block. Each block is designed to include tasks addressing different functions.

Question processing component. To generate an answer, the linguistic analysis of the question has to be performed in order to identify the syntactic and semantic structure. The question can be entered by the user as text or through vocal interaction. In the second case, besides text analysis, the question processing involves phonetic analysis and prosodic features extraction.

Answer generation component. The answer generation process is designed to take as input the output of the question processing component (i.e. text analysis and, if available, prosodic features). In this phase we classify the question, disambiguate the question meaning and infer some characteristics of the user. The goal is to provide a personalized response, which in our model is a learning path containing all the relevant concepts for answering the user's question. The PLP is retrieved from an ECM which contains all the concepts, their relations and associated materials. We select which concepts to retrieve and show to the user according to the user's request, the features extracted and the topology of the map. Details on these tasks are discussed in the next sections.

3 QUESTION TYPE CLASSIFICATION

Question Classification is the task of assigning a question to its corresponding category. It can be performed by matching, through heuristics, the syntactic structure of the question with manually defined patterns [24] or adopting machine learning techniques trained on hand-labeled databases of questions [26]. The first method suffers a drop of accuracy when new question patterns are encountered, but works very efficiently on predefined questions. Machine learning strategies are more adaptive but require large annotated datasets. Sometimes hybrid approaches are adopted [5, 14]. We adopt a rule-based approach where the classification is performed considering the output of the text analysis, as shown in Figure 1. The classification can be conducted with regards to the focus (e.g. the word, or sequence of words, that identifies what the question is searching for [19]) and the formal linguistic structure of the question. We handle three question types, following the classification in [18], with specific rules as follows:

- **Factoid questions (FQ)**: WH- questions (excluding "why" and "how" questions) where answering involves a single short phrase or a sentence. In our case, we redefine FQ as questions that refer to specific concepts or their properties in line with providing the answer as PLP. Although FQ are a typical form of questions, their superficial structure can vary a lot. With the text analysis we can identify questions' core elements and structures, for example the presence of coordinated constructions, also distinguishing between conjuncted items or clauses. For our purposes, we further distinguish three sub-types of FQ:
  - *Definition questions* ask about a specific concept or about a concept property. A typical wording of definition question is *"What is topic_X?"*.
  - *Comparison questions* ask about differences and similarities between two or more topics. Comparison questions are usually expressed as *"What are the differences between topic_X and topic_Y, and topic_N?"*.
  - *List questions* are a special case of FQ, where the answer concerns specific instances or properties of a topic, such as *"What are the types of topic_X?"*.
- **Causal questions**: "how" and "why" questions, used to ask for clarifications, explanations, reasons and elaborations.

They can refer to properties of a topic or to relations between topics. Text analysis here is fundamental to be able to make assumptions about the user's real intention, because causal questions can be ambiguous: a question like *"Why the topic_X has a property_Y?"* can be interpreted as *"Why the topic_X and not topic_Z?"* or *"Why it has property_Y?"*.

- **Confirmation questions**: the user asks if a certain notion is true or false. We recognize two types of confirmation questions: questions referring to specific properties of a concept (i.e. *Does property_Y belong to topic_X?*), and those referring to the relation between concepts (i.e. *Are topic_X and topic_Y related? Is topic_X a prerequisite of topic_Y?*).

Text analysis for question classification. In order to extract the information for question type classification and focus identification, the model proposes to perform question processing in the textual analysis phase. For this task well-established techniques are available. Question processing involves a series of pre-processing tasks, like morpho-syntactic analysis, Named-Entity tagging and predicate-argument structure extraction, like in [14]. Based on the pre-processed text, we perform keyword extraction and formal patterns identification. Keywords are used to acquire the question focus, while formal patterns, referring both to lexical and syntactic analysis, are for question classification and focus extraction. The most informative formal patterns that we identified are the following:
- *Predicate structure*: Semantic role labeling can be obtained through semantic shallow parsing. Semantic roles help distinguishing the meaning of a predicate and recognizing semantically similar sentences (the syntactic structure may change but the semantic annotation remains the same). Like in previous studies (see [14, 16, 23]), semantic roles can be used to enhance the question focus with its taxonomic category. On the surface level, the question is analyzed considering syntactic roles (subject, object, adjuncts, etc.) and their dependency relations, which are utilized to recognize specific construction and content words.
- *WH- headword*: WH-words are interrogatives that generally introduce a question. Possible WH-words are *what, which, where,* etc. The complete absence of WH-words can suggest the presence of a confirmation question or a request like *Name all the properties of topic_X*). In those cases we follow the methodology described in [12] and only look at the syntactic relations in order to identify the head word of the question, that is the word specifying the object of the request.
- *Causal construction*: Causal questions are mostly introduced by *why* or *how*, but their main characteristic is that they put two or more elements into a causal relation. [11] identifies some lexical choices and syntactic relations that are strong cues of the presence of a causal construction.
- *Confirmation-seeking constructions*: They appear when asking if a certain fact is true of false. *Is it true that* and similar wordings are strong triggers of confirmation questions, so we treat them like WH- words.
- *Belongings*: In case of questions referring to specific properties of a topic, usually the predicate expresses belonging. If not, it might be that the property is a verb (like in the example *Does the fire burn?*, *burn* refers to the property of burning), and those terms are strictly related to the domain of interest.

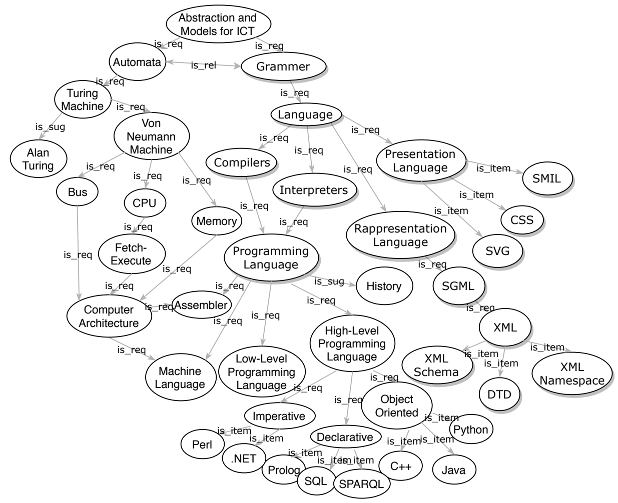

Figure 2: ECM example

Sometimes questions, both oral and typed, can be incomplete, ambiguous, misspelled or syntactically incorrect. We attest those characteristics when analyzing the question, but we do not take them into account when trying to classify the question type or extract the question focus. However, they are considered when extracting the user competency level and confidence with the topics.

4 ANSWER BUILDING

The core task of the answer generation component is the ECM Pruning. The returned ECM will be used for retrieving the PLP that is presented to the user as the answer to her/his question. As shown in Figure 1, ECM Pruning takes as input:

- the question type (from Question Classification)
- the question focus (from Focus Extraction), and
- speech-based features (from Prosody Analysis)

Question Classification was presented in the previous section. Focus extraction and prosody analysis are briefly discussed below.

Focus Extraction is the identification of the ECM concept(s) mentioned in the question. Three possible cases are handled in the following order:
1. Direct Match: The keyword(s) or the named entity(ies) in the question perfectly correspond to ECM concepts.
2. Indirect Match: If no direct match exists, query expansion is performed by exploiting external knowledge (e.g., Framester [10]). We try to match ECM concepts with WordNet synsets related to the keyword term(s) or DBpedia entities referring to Named Entity(ies) found in the question, similarly to [22].
3. Inferred match: If no indirect match is possible, a further option is to try a match with the learning materials associated to each ECM concept by exploiting Information Retrieval (IR) techniques.

Prosody Analysis and feature extraction. Our model allows for vocal interaction, which allows us to collect a variety of information about the level of confidence with the topics and the user's emotional state. This information can be obtained from the acoustic prosodic analysis of the speech [7, 15, 21]. Using a tool like the one described in [4], we can obtain a phonetic annotation of the transcribed user request and observe the prosodic aspects of the interaction. Like a teacher who listens to a question, including its prosody and emotional aspects, and decides what learning material best fits the learner's needs, our aim is to identify some features which could be an indicator of her/his confidence with the topic and engagement. D'Mello's definition of engagement in terms of "goal-directed state of active and focused involvement in a learning activity" [6] perfectly fits our objective, that is to identify cues about the willingness of the learner to get more knowledge about a specific topic. [6] Following [8, 9, 17], we assume that speech rate, pitches and high tone are an expression of higher engagement, while long pauses indicate uncertainties and a low level of confidence with the topic. In any case, we plan to make investigations about the applicability of results from the literature to our context, considering specifically short vocal questions.

4.1 ECM Knowledge Base

This section reports some aspects of the formal definition of the ECM model [1] underlying the knowledge base and the PLP generation. An ECM represents a general structure of the subject matter domain that explicits the concepts, their learning materials and the hierarchical and associative relations that connect the concepts and forms the ECM. The concepts can be of type {primary notion, secondary notion, learning outcome}, while the educational relations can be of type {is_req, is_item, is_sug, is_rel}. Is_req (is_requirement) is a propaedeutic precedence relation, is_item defines hierarchical hyponym/hypernym or meronym structures, is_sug (is_suggested) relates a main concept with its in-depth examination and the is_rel (is_related) represents a relation between closely related concepts without prerequisite constrains. By means of the ECM and its algorithm for learning path generation, it is possible to find a "suitable" teaching and learning path through the pruned concepts of the ECM.

4.2 Pruning and PLP generation

The process of pruning aims to identify the ECM concepts and arcs that satisfy the question focus and the learner's educational needs (as sketched above). We call *unit of learning* (UoL) the output of the pruning task. The result is still a concept map, which then has to be linearized into a sequence of concepts in a specific order. With respect to the different **types of focus-map matching**, we distinguish the following pruning strategies:

- **Direct Match Pruning** is further split into three cases:

1. *Definition questions*: In case the focus is a single concept in the ECM, e.g., "What is a Turing Machine", the basic pruning rule is: this(<TOPIC_X>) {return:<MATERIAL>}. In the more complex case, e.g. "What is the fetch-execute cycle of CPU?", in which the focus is fetch-execute cycle but also CPU is identified, the pruning rule first searches for a path in the ECM graph between the two concepts. If the path is found, then it is returned as pruned map, otherwise the returned LP is composed of the two paths connecting each of the identified concepts to their closer common prerequisite.
2. *Comparative questions*: If TOPIC_X and TOPIC_Y have a common prerequisite and learning outcome and if the distances between the common prerequisite of TOPIC_X, TOPIC_Y and between TOPIC_X, TOPIC_Y and their learning outcome are below a threshold, then it returns the UoL starting with the common prerequisite and ending with the learning outcome. Example: "What is the difference between compiler and interpreter?" returns the map containing {Language, Compiler, Interpreters, Programming Language}. Else, returns the direct_req for both the concepts (see Figure 2).
3. *List questions*: The pruning is performed by including the outgoing arcs (is_item or is_req) of the concept identified as focus. In case there are no is_item relations, then a further IR analysis is performed on the content of the subsidiary concepts, in order to identify if those concepts are in hypernym/hyponym or meronym relation with the prerequisite concept. For instance "What kind of High-Level Languages does exist?" results in the map containing {Imperative, Declarative, Object Oriented}. This was the case when the question has single focus concept. Otherwise, in cases such as "What are the types of declarative and imperative languages?", the pruning should contain a LP starting from the first common prerequisite of both concepts (if path_distance(focus concept, common prerequisite)<threshold), thus the map will contain {High-Level Language, Declarative, Prolog, SPARQL, SQL, Imperative, Perl, Python}.
4. *Causal questions*: In case the concepts involved in the causal question are reachable from one another in the map (there exists a path that connects them), then the response to the learner question is the path itself. For instance "What makes Machine Language, CPU dependent?", response {CPU, fetch-execute, Computer Architecture}.
5. *Confirmation questions*: Once the focus is identified on the map, an IR search of lexical pattern on its learning material, and on its neighbor (adjacent) concepts, can be executed. Example: "Does an interpreter language execute the program directly?".

- **Indirect Match**. If concepts are found that match the ECM, the pruning process includes the concepts found plus the concepts that are on the path in between the extracted concepts. As a result, the new pruned map becomes the map on which the above rules of Direct Match are applied.

- **Inferred Match**. When neither of the aforementioned cases resolve to ECM focus mapping, the search is performed with IR techniques on the information layer of the ECM, by applying lexical-patterns defined on the base of the type of question (see Section 3). Example: "How does the computer work?" since there is no direct match for computer and nor indirect match (after the External Knowledge enrichment), then a bottom-up search is performed on the information layer of the ECM. The result of this retrieval is the identification of the following concepts {Von Neumann Machine, Bus, CPU, Memory, Fetch-Execute, Computer Architecture}.

The above pruning rules can be further refined by considering the **Confidence and Engagement level**, when these features are available (extracted from the prosody analysis). In our model the confidence level with a topic is used to define how much prerequisite or subsidiary knowledge should be presented to the learner, while the engagement impacts the depth of knowledge to be provided. The methods have the following forms: direct_req(TOPIC_X) returns all the concepts that are direct prerequisite of TOPIC_X; this(TOPIC_X) returns the current concept; for_every_UoL_topic returns all the concepts that are prerequisite of the focus concept till the UoL root concept. For instance: (i) high confidence results in less prerequisite knowledge needed, and more direct and subsidiary knowledge for the requested concept(s), moreover, in case also high engagement is detected, then further concepts strongly related (is_item, is_rel, is_sug) to the requested concept are presented, i.e. this(<TOPIC_X>) return{is_item*, is_sug*,is_rel*, is_req* <MATERIAL>}; (ii) instead, in case the analysis returns high confidence while engagement is not detected, the pruning contains the focus concept and its subsidiary (adjacent) concepts to which <TOPIC_X> is prerequisite, i.e. this(<TOPIC_X>) return{ is_req*, <MATERIAL>}.

The final **PLP generation** is the topological sort of the concepts in the UoL resolved in a graph. This linearization preserves the precedence of the prerequisites defined by the is_req relationship. E.g., if TOPIC_X, representing the concept "Automata", is a propaedeutic requirement of the TOPIC_Y, representing the concept "Turing Machine", in the map exists an association of type is_req between the two concepts (nodes); instead, in the linearized sequence this means that TOPIC_X precedes TOPIC_Y [1].

5 DISCUSSION AND RELATED WORK

We presented a model for question answering specifically designed for educational interactive environments. The main contribution of our approach is that it organizes the answer as a learning content, which is tailored according to the question type and focus and, when possible, the learner features estimated from the vocal question (confidence with the topic and engagement).

The approach could be well-integrated into MOOC platforms that use interactivity. There are several examples of MOOCs that implement such features. MOOCBuddy is a chatbot based on Facebook Messenger Platform acting as a MOOC Recommender System that uses user's social media profile and interests to achieve personalization of the learning experience [13]. Automatic Speech Recognition can be embedded in such systems so prosodic information can also be acquired from the interaction, whether it is a dialogue or a single question from the user. To achieve this goal, there are tools that perform the transcription of vocal interaction into text, together with the automatic annotation of prosodic elements [4]. Prosody can provide information about a speaker's mood, helping the recommender system to take decisions accordingly (for a survey about emotion recognition using speech see [3]). Future work will address the investigation of real possibilities and conditions to estimate confidence level and engagement from short vocal interactions (questions) and improve the personalization accordingly. Also, since users interacting with vocal systems use free-style speech, further work is needed to identify the question focus and type for cases that are not currently addressed by the model.

REFERENCES

[1] Giovanni Adorni and Frosina Koceva. 2016. Educational concept maps for personalized learning path generation. In *Conference of the Italian Association for Artificial Intelligence*. Springer, 135–148.

[2] Thushari Atapattu, Katrina Falkner, and Nickolas Falkner. 2015. Educational question answering motivated by question-specific concept maps. In *International Conference on Artificial Intelligence in Education*. Springer, 13–22.

[3] S. Basu, J. Chakraborty, A. Bag, and M. Aftabuddin. 2017. A review on emotion recognition using speech. In *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*. 109–114. https://doi.org/10.1109/ICICCT.2017.7975169

[4] Brigitte Bigi and Daniel Hirst. 2012. SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody. In *Speech Prosody*. 1–4.

[5] Caner Derici, Kerem ÃĞelik, Ekrem Kutbay, YiÄ§it AydÄśn, Tunga GÃijngÃűr, Arzucan ÃŰzgÃijr, and GÃijnizi Kartal. 2015. Question Analysis for a Closed Domain Question Answering System. In *Computational Linguistics and Intelligent Text Processing*, Alexander Gelbukh (Ed.). Vol. 9042. Springer International Publishing, Cham, 468–482. https://doi.org/10.1007/978-3-319-18117-2_35

[6] Sidney D'Mello, Ed Dieterle, and Angela Duckworth. 2017. Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational psychologist* 52, 2 (2017), 104–123.

[7] Sidney K D'mello, Scotty D Craig, Amy Witherspoon, Bethany Mcdaniel, and Arthur Graesser. 2008. Automatic detection of learnerâĂŹs affect from conversational cues. *User modeling and user-adapted interaction* 18, 1-2 (2008), 45–80.

[8] Harishchandra Dubey, Abhijeet Sangwan, and John HL Hansen. 2017. Using speech technology for quantifying behavioral characteristics in peer-led team learning sessions. *Computer Speech & Language* 46 (2017), 343–366.

[9] Kate Forbes-Riley and Diane Litman. 2011. Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication* 53, 9-10 (2011), 1115–1136.

[10] Aldo Gangemi, Mehwish Alam, Luigi Asprino, Valentina Presutti, and Diego Reforgiato Recupero. 2016. Framester: a wide coverage linguistic linked data hub. In *European Knowledge Acquisition Workshop*. Springer, 239–254.

[11] Roxana Girju. 2003. Automatic detection of causal relations for question answering. In *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering-Volume 12*. ACL, 76–83.

[12] Zhiheng Huang, Marcus Thint, and Zengchang Qin. 2008. Question classification using head words and their hypernyms. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 927–936.

[13] Adrian Iftene and Jean Vanderdonckt. 2016. Moocbuddy: a chatbot for personalized learning with moocs. In *RoCHIâĂŞInternational Conference on Human-Computer Interaction*. 91.

[14] Adam Lally, John M. Prager, Michael C. McCord, Branimir K. Boguraev, Siddharth Patwardhan, James Fan, Paul Fodor, and Jennifer Chu-Carroll. 2012. Question analysis: How Watson reads a clue. *IBM Journal of Research and Development* 56, 3.4 (2012), 2: 1–2: 14.

[15] Chi-Chun Lee, Jangwon Kim, Angeliki Metallinou, Carlos Busso, Sungbok Lee, and Shrikanth S Narayanan. 2014. Speech in affective computing. In *The Oxford Handbook of Affective Computing*.

[16] Xin Li and Dan Roth. 2006. Learning question classifiers: the role of semantic information. *Natural Language Engineering* 12, 3 (2006), 229–249.

[17] Yuanchao Li, Carlos Toshinori Ishi, Nigel Ward, Koji Inoue, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara. 2017. Emotion Recognition by Combining Prosody and Sentiment Analysis for Expressing Reactive Emotion by Humanoid Robot. In *Proc. of APSIPA Summit and Conference*, Vol. 2017. 12–15.

[18] Amit Mishra and Sanjay Kumar Jain. 2016. A survey on question answering systems with classification. *Journal of King Saud University-Computer and Information Sciences* 28, 3 (2016), 345–361. https://doi.org/10.1016/j.jksuci.2014.10.007

[19] Dan Moldovan, Sanda Harabagiu, Marius Pasca, Rada Mihalcea, Roxana Girju, Richard Goodrum, and Vasile Rus. 2000. The structure and performance of an open-domain question answering system. In *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*. ACL, 563–570.

[20] Sandra Sanchez-Gordon and Sergio Luján-Mora. 2016. How Could MOOCs Become Accessible? The Case of edX and the Future of Inclusive Online Learning. *Journal of Universal Computer Science* 22, 1 (2016), 55–81.

[21] Björn Schuller, Anton Batliner, Stefan Steidl, and Dino Seppi. 2011. Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication* 53, 9-10 (2011), 1062–1087.

[22] K. Sengloiluean, N. Arch-int, S. Arch-int, and T. Thongkrau. 2017. A semantic approach for question answering using DBpedia and WordNet. In *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. 1–6. https://doi.org/10.1109/JCSSE.2017.8025918

[23] Dan Shen and Mirella Lapata. 2007. Using semantic roles to improve question answering. In *Proc. of the joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*.

[24] JoÃčo Silva, LuÃŋsa Coheur, Ana Cristina Mendes, and Andreas Wichert. 2011. From symbolic to sub-symbolic information in question classification. *Artificial Intelligence Review* 35, 2 (Feb. 2011), 137–154. https://doi.org/10.1007/s10462-010-9188-4

[25] Ayse Saliha Sunar, Nor Aniza Abdullah, Su White, and Hugh C Davis. 2015. Personalisation of MOOCs. In *Proceedings of the 7th International Conference on Computer Supported Education-Volume 1*. SCITEPRESS-Science and Technology Publications, Lda, 88–97.

[26] Nguyen Van-Tu and Le Anh-Cuong. 2016. Improving question classification by feature extraction and selection. *Indian Journal of Science and Technology* 9, 17 (2016).