INTEGRA: A Web-Based Differential Diagnosis System Combining Multiple Knowledge Bases

Aris Papakonstantinou
Department of Electrical and Computer Engineering
Hellenic Mediterranean University,
Heraklion, Crete, Greece
aris.papakonstantinou@outlook.com

Haridimos Kondylakis
Institute of Computer Science, FORTH,
Heraklion, Crete, Greece
kondylak@ics.forth.gr

Emmanouil Marakakis
Department of Electrical and Computer Engineering,
Hellenic Mediterranean University,
Heraklion, Crete, Greece
mmarak@cs.teicrete.gr

ABSTRACT
Diagnosis is one of the most important tasks of health providers with a tremendous impact on patient's health and well-being. Although many diagnostic decision support systems have been proposed so far, with several advantages and disadvantages, we have not yet seen approaches trying to integrate the decision of individual systems to improve the quality of the final diagnosis. To this direction, we present INTEGRA a novel decision support system, allowing multiple underlying diagnostic decision-support sub-systems to work in parallel, effectively integrating the individual decisions, increasing as such the quality of the final diagnosis. In addition, INTEGRA enables health providers to get explanation of the decisions based on the combined diagnosis and further explore the recommendations of the individual decision support sub-systems.

CCS CONCEPTS
Information systems → Expert systems

KEYWORDS
Expert Systems, Differential Diagnosis, Decision Support, Meta-rules

1. INTRODUCTION
Diagnosis is one of the most important tasks performed by health providers and the impact of diagnostic errors on patient safety has been highly recognized [1]. However, although quantifiable (for example in the Harvard Medical Practice Study, diagnostic errors accounted for 17% of preventable errors [2]) limited attention has been shown at improving diagnostic errors and thousands of patients die or suffer every year due to them.

ICT technology has the potential to actually empower both patients and health providers [3]. To this direction, *expert systems* could benefit both the diagnostic procedure and the education of health providers. An expert system is a computer system that emulates the decision-making ability of a human expert. The development of such a system requires the relevant knowledge to be extracted from an expert and then to be represented in knowledge base (KB) for reasoning [4]. Based on this knowledge, it emulates the decision-making ability of a human expert and can aid human experts in decision making and in education of medical personnel.

In this paper, we present INTEGRA, a novel decision support system integrating multiple diagnostic sub-systems combining their results for optimizing the result diagnosis. In addition, our system is able to visualize the explanations for the suggested decisions both at the integrated decision level and at the decisions of the individual underlying sub-systems. Preliminary evaluation shows the added value of our system, showing an increase on the number of the correct diagnoses on respiratory cases.

A preliminary version of the system was developed for educational purposes in medical schools, supporting only a single KB [5]. The version reported here, is not only for educating medical personnel but for using it as a decision support tool in the diagnosis process. In addition, it now supports not only multiple underlying diagnostic sub-systems but also multiple knowledge bases. These knowledge bases can be used for independent diagnosis and their diagnosis results are combined. The multiple diagnostic sub-systems and knowledge bases are transparent to the user who gets a list of possible diseases sorted in ascending order of the derived likelihood. This feature significantly improves the quality of the proposed diagnosis.

The remaining of this paper is structured as follows: In Section 2, we elaborate on related work. Then, in Section 3, we present the system architecture and we elaborate on the various components and algorithms used. In Section 4 we present a preliminary evaluation of our work on respiratory cases and finally, Section 5 concludes this paper and presents directions for further work.

PETRA '20, June 30-July 3, 2020, Corfu, Greece

ACM ISBN 978-1-4503-7773-7/20/06…$15.00
https://doi.org/10.1145/3389189.3397980

2. RELATED WORK

Within the years, several individual expert systems have been developed for decision support on diagnostic procedures.

APACHE III for example is a system [6] trying to predict the person's risk of dying in a hospital. This prediction is based on a comparison of the medical history of 18,000 cases, stored in the system database. It has an average of 95% predictive accuracy, and it uses a score-based mechanism. LISA on the other hand, is a Clinical Information and Decision Support System for co-operative care in childhood acute lymphoblastic leukaemia [7]. It is primarily concerned with providing support during the patient's treatment period, where weekly decisions on drug dosing should be made. Dose adjustment rules are applied using the Guideline Modelling Language PROforma and the Recommendations are provided in clinical setting by the TALLIS PROforma approval mechanism. ISABEL is a web-based clinical decision support system providing support for paediatric diagnostic decisions [8]. Isabel uses Autonomy's natural language processing software and consists of a proprietary medical database with over 11,000 diagnoses and 4,000 drugs. It supports templates for querying and supports Health Level Seven (HL7) for interoperability. PUFF is an expert system introduced to interpret pulmonary function test data [9]. Its reasoning is based on a backward chaining and it uses about 400 rules in knowledge base.

Therapy Edge HIV is a web-based clinical decision support system introduced in 2005 that deals with HIV treatment [10]. Its reasoning is based on temporal guidelines to assess the patient’s current state and create alternative treatment options. Therapy Edge HIV implements an API to communicate with external systems, providing information in XML format. The expert system in [11], [12] derives differential diagnosis of epilepsy in childhood, using meta-rules. The approach in these two works has one meta-rule with 28 diagnostic criteria. These 28 diagnostic criteria constitute the terms of the precondition of the meta-rule and they are connected with the $\wedge$ (and) operator. Instances of this meta-rule are derived during the diagnosis process. The precondition of each instance consists of $\wedge$-connected instances of terms of the meta-rule. That is, each instance of the meta-rule may have some of the diagnostic criteria. The summation of the weights of the terms of the precondition of each instance has to be in the range [0, 1]. The summation of the weights at the end of diagnosis is mapped to a percentage. This is the expert’s confidence for the truth of the derived type of epilepsy for a specific patient. Docs ‘n Drugs [16] is an intelligent tutoring system for web-based and case-oriented training in medicine, however is was identified to have a poor user experience. There, the development of a training case influences the correctness of the learner's answers, whereas ICD-10 and other ontologies are also used.

We believe using ontologies [17] is to the right direction, and especially exploiting ICD-10 is a key as it is widely used in health and medical systems around the world. In addition, user experience in the aforementioned systems is low. Although we have explored in the past approaches dealing with uncertainty [18], [19], in the current approach, we would like to investigate a simpler, cleaner approach. Finally, to the best of our knowledge INTEGRA is the first system supporting multiple knowledge bases, offering both individual KB explanation of the results and an integrated decision outcome. Our system has been designed to be extensible, enabling the uninterrupted addition of new diagnostic sub-systems at the GUI, the Diagnostic or the KB level as we shall see in the sequel.

3. ARCHITECTURE

INTEGRA employs a three-layered architecture, shown in Figure 1. It consists of the web interface, the diagnostic sub-system and the knowledge base layer. However, multiple individual sub-systems can co-exist in these three layers, eventually integrated at the GUI layer, offering different aspects on the decisions proposed. In the sequel, we describe in details each one of these layers.

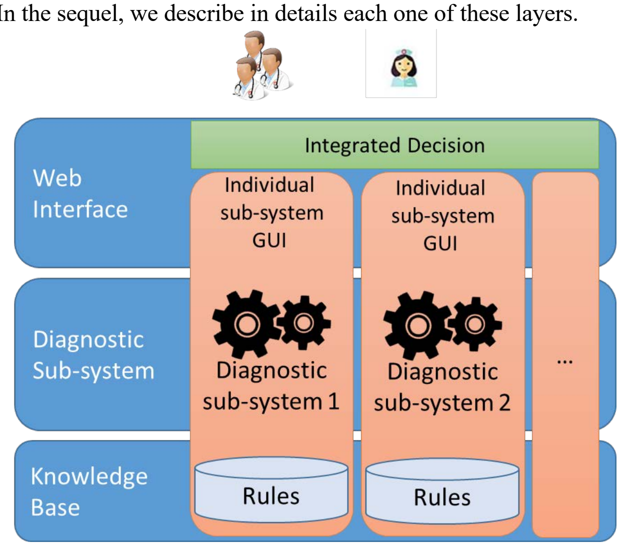


Figure 1: INTEGRA architecture.

3.1 The Graphical User Interface

This layer has been implemented using CSS, Javascript and HTML. The expert can enter patient’s common complaints and select the corresponding symptoms from a generated list. A screenshot from the system, presenting those lists is shown in Figure 2.

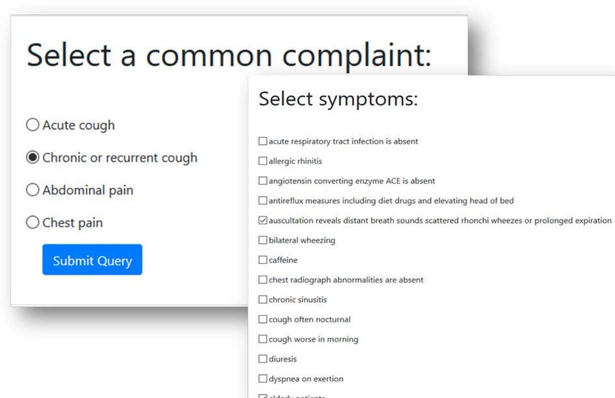
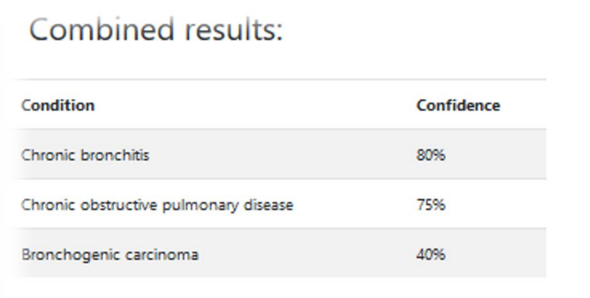


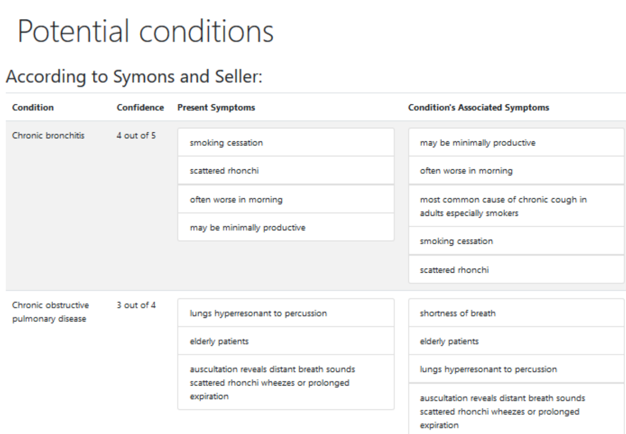
Figure 2. Selecting a common complaint and symptoms.

As soon as the expert makes the appropriate selections, the results from the diagnosis are presented along with a confidence likelihood. A screenshot of the presented results is shown in Figure 3. As shown, the combined results are presented along with a confidence score. In addition, the user can see the differential diagnosis based on the individual knowledge base and diagnostic sub-system used. For example, the diagnosis shown in Figure 4 is proposed according to the first diagnostic sub-system which has been constructed based on Symons and Sellers textbook [22]


Combined results:

Condition	Confidence
Chronic bronchitis	80%
Chronic obstructive pulmonary disease	75%
Bronchogenic carcinoma	40%

Figure 3. Suggesging potential diagnosis.


Potential conditions

According to Symons and Seller:

Condition	Confidence	Present Symptoms	Condition's Associated Symptoms
Chronic bronchitis	4 out of 5	smoking cessation scattered rhonchi often worse in morning may be minimally productive	may be minimally productive often worse in morning most common cause of chronic cough in adults especially smokers smoking cessation scattered rhonchi
Chronic obstructive pulmonary disease	3 out of 4	lungs hyperresonant to percussion elderly patients auscultation reveals distant breath sounds scattered rhonchi wheezes or prolonged expiration	shortness of breath elderly patients lungs hyperresonant to percussion auscultation reveals distant breath sounds scattered rhonchi wheezes or prolonged expiration

Figure 4. Diagnosis according to Symons and Seller and justification.

Besides presenting the diagnosis according to an individual sub-system an explanation is also shown. The example shown in Figure 4 further explains why Chronic Bronchitis is suggested - as four out of the five potential symptoms are present in the current situation. In addition, the system presents other associated symptoms with the specific case.

3.1.1 Integrating Decisions

Using our system, many diagnostic machines can work in parallel. To this direction, the expert selections for a specific patient are fed simultaneous to all underlying diagnostic sub-systems. The derived diagnosis from all sub-systems is then combined and shown to the user in a way that the individual diagnoses are transparent. That is, the diagnosis of possible diseases is listed in increasing order of their likelihood. The combination of the diagnosis of the individual diagnostic sub-systems is performed as follows:

1. The diagnostic results returned by a single diagnostic sub-system are placed in a list ordered by their confidence score.
2. The confidence of the diagnostic results returned by more than one diagnostic sub-systems is the largest confidence returned by the individual machines.

Then, the corresponding results are visualized, they are ordered by their confidence to the end-user as shown in Figure 3. Although for combining the confidence of the individual diagnostic results for the evaluation of this paper we rely on the max confidence, we are also exploring "aggregation designs" based on min, max, Borda, and Fair [13], [14], [15] which we intend to report in the journal version of this article.

3.2 The diagnostic sub-systems

In this layer we have the individual diagnostic sub-systems. The diagnostic sub-systems can either exploit a single KB already available, provide their own KB or combine multiple KBs. The only restriction is to provide the appropriate APIs and GUIs for their effective integration to the whole platform.

In our case, two individual diagnostic sub-systems are used. The first one is an instance of the system presented in [5] which is based on the medical diagnostic reasoning in [22], whereas the second is based on the medical diagnostic reasoning in [21], each one fed with a different KB as we will describe it in the next section. The approach in [20] has been used for the derivation of likelihoods in both diagnostic sub-systems.

For reasons of completeness, we just note that in [5], the reasoning process is driven using a meta-rule, considering complaints, nature of patient (conditions most relevant with a particular subgroup), nature of symptoms (conditions amplifying the symptoms), associated symptoms, precipitating and aggravating factors, ameliorating factors, physical findings and diagnostic studies. During the diagnosis process instances of this meta-rule are created. One instance of the meta-rule is created during each diagnostic task. Each derived instance is based on different prime complain from patient. Examples of prime complaints are headache, sore throat, menstrual pain etc.

The reasoning of the other diagnostic sub-system is based in [21]. It is built on the same concept, that is, its KB consists of one meta-rule. Again, an instance of the meta-rule is derived based on the set of patient's symptoms. There is not the idea of prime complaint (symptom) in the second diagnostic sub-system as it is in the first sub-system. Certainly, some symptoms are more important than other for the final decision. The derivation of final diagnosis does not make any discrimination between symptoms and the final decision is based evenly on the overall set of symptoms. The execution of derived rule, instance of meta-rule, performs the diagnosis of the particular case of symptoms. Both diagnostic sub-systems rely on the notion of likelihood ratios (LRs) as diagnostic weights [20]. The likelihood ratio (LR) of a physical sign is defined as the proportion of patients *with* disease who have a particular finding divided by the proportion of patients *without* disease who also have the same finding [20]. According to the idea of likelihood

Condition	Nature of Patient	Nature of Symptoms	Associated Symptoms	Precipitating and Aggravating Factors	Ameliorating Factors	Physical Findings	Diagnostic Studies
Chronic or Recurrent Cough							
Postnasal drip	May not be aware of condition	Frequent throat clearing and hawking. Cough worse in morning		Recumbency Chronic sinutis Vasomotor rhinitis Allergic rhinitis Nonallergic rhinitis with eosinophilia		Mucoid secretions in posterior pharynx Palpation, percusssion, and transillumination of sinuses reveal sinusitis Mucosa of nose/oropharynx:cobbmestone	
Asthma	May have family history of allergies, atopy, or asthma	Recurrent cough Minimally or not productive (if productive, secretions clear and mucoid)	Shortness of breath	Exercise May be worse during seasonal allergies		Bilateral wheezing	Pulmonary function tests Response to isoproterenol and methaclorine
Chronic bronchitis	Most common cause of chronic cough in adults (especially smokers)	May be minimally productive Often worse in morning			Smoking cessation	Scattered rhonchi	
Chronic obstructive pulmonary disease	Elderly patients		Shortness of breath			Lungs hyperresonant to percussion Auscultation reveals distant breath sounds, scattered rhonchi wheezed, or prolonged expiration	Pulmonary function tests

Figure 5: Differential diagnosis of cough according to [1]

ratios, if a physical sign characteristic of a suspected diagnosis is *present* (i.e., positive finding), that diagnosis becomes more likely.

If the characteristic finding is *absent* (i.e., negative finding), the suspected diagnosis becomes less likely. How much these positive and negative results modify likelihood, however, is distinct for each physical sign. Some findings, when positive, shift likelihood upward greatly, but they change it little when negative. Other signs are more useful if they are absent because the negative finding practically excludes disease, although the positive one changes likelihood very little. More details on the evaluation of likelihood of a disease in [20].

3.2 The Knowledge Base Layer

In this layer, we can find the individual KBs used by the various diagnostic sub-systems on top. We choose to exploit common existing knowledge on differential diagnosis, taught to medical schools to this purpose.
As such, we designed two different meta-rules based on two well-known approaches which are presented [21] and [22] respectively. The constructed meta-rules have been encoded as Prolog meta-rules. These meta-rules are part of KB of each sub-system. The other part of the KB are the rules which are derived dynamically

based on patient's symptoms. The current state of INTEGRA is in prototype form because of that the KB of each sub-system can perform diagnosis on some sets of symptoms. In order to be able to discuss the design idea of our system and evaluate our prototype system both diagnostic sub-systems can process a subset of corresponding symptoms. For example, the first diagnostic sub-system performs full diagnosis for cases of patients with prime symptom "cough". Moreover, in the second diagnostic sub-system, "cough" is included in all sets of symptoms that it can perform diagnosis. An example from [22] for diagnoses related to cough is shown in Figure 5. For more information on the internal representation of meta-rules, on their instances and on the performed reasoning the interested reader is forwarded to [5].

4. EVALUATION

To evaluate our system, we performed a preliminary analysis based on respiratory problems. Initially, we designed the corresponding meta-rules based on medical diagnosis discussed in [21] and in [22] for each diagnostic sub-system and then we also identified respiratory clinical cases for which we had both the symptoms and the confirmed diagnosis by medical experts. Those cases were extracted from [23] and [24] – unfortunately there were only twenty respiratory cases.

The results are shown in Figure 6, where we identify cases that the two diagnostic sub-systems and the integrated decision were correct (marked as 1 in the table). As identified, individual diagnostic sub-systems are able to diagnose a condition if the condition is part of the knowledge base provided by the expert. The second diagnostic sub-system has a lower diagnostic score than the first machine, 30% versus 55% of the total cases, whereas the integrated decision is able to correctly diagnose the 60% of the cases. By adding new knowledge to the meta-rules of each sub-system it is possible to reach higher diagnostic percentage.

As a next step we intend to extend the meta-rules in the two sub-systems in the following directions. The first sub-system will be extended from medical knowledge in order to be able to perform diagnosis for all set of prime symptoms discussed in [22]. The second sub-system will be extended by medical knowledge in the same direction to be able to process corresponding sets of symptoms. In addition, we will enhance the meta-rules in both sub-system with medical knowledge which is not available in [22] and in [21] and we consider it is important for the diagnosis process. Then, eventually we will be able to evaluate our system based on the whole list of 350 cases that can be extracted from [23] and [24].

5. CONCLUSIONS AND DISCUSSION

In this paper we present INTEGRA, a novel solution enabling the combination of multiple KBs and decision support sub-systems for enhancing the quality of the diagnosis. Our solution supports explanation of the results from the individual sub-systems and is extensible and self-intuitive, based on accumulated validated medical knowledge used to teach medical experts. To the best of our knowledge we are the first to support differential diagnosis combining multiple knowledge bases.

Condition to test	First Sub-system	Second Sub-system	INTEGRA
Pleural effusion			
Pleural rub			
Asthma	1	1	1
Chronic bronchitis	1		1
Bronchiectasis			
Cor pulmonale	1		1
Consolidation			
Bronchogenic carcinoma		1	1
Cystic fibrosis			
Fibrosing alveolitis	1		1
Pulmonary fibrosis	1		1
Pneumothorax			
Old tuberculosis	1		1
Pickwickian syndrome			
Collapsed lung			
Bacterial Pneumonia	1	1	1
Mycoplasmal bronchitis	1	1	1
Postnasal drip	1	1	1
Asthma	1	1	1
COPD	1		1
	11	6	12
Percentage	**55**	**30**	**60**

Figure 6: Preliminary evaluation results

As immediate next step is to generalize our evaluation on the whole list of the 350 cases already identified, extending the meta-rules in both directions as discussed in the previous section. Then, further integration designs will be explored besides maximum, such as minimum, average, Borda, fair etc.

In addition, an interesting direction would be to use machine learning for combining the outcomes of the individual diagnostic sub-systems and check whether using machine learning techniques could further improve the quality of the generated results. We could also add a new diagnostic sub-system which will perform diagnosis based on machine learning techniques.

Further, the derived instances of meta-rules, i.e. individual rules, and symptoms might be many a visual summary [25], [26], [27] could be useful for quickly guiding users on the decisions proposed.

ACKNOWLEDGEMENT

Work presented in the paper is part of the BOUNCE project and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777167. Any opinions, results, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of BOUNCE or the European Commission.

REFERENCES

[1] H. Singh, L.A Petersen, E.J. Thomas (2006). Understanding diagnostic errors in medicine: a lesson from aviation. *Quality & Safety in Health Care*, 5(3), 159-164.

[2] L.L. Leape, T.A. Brennan, M.P.H. Nan Laird, et al. (1991). The Nature of Adverse Events in Hospitalized Patients - Results of the Harvard Medical Practice Study II. N Engl J Med., 324(6):377-84.

[3] H. Kondylakis, A. Bucur, C. Crico, et al. (2020). Patient Empowerment for Cancer Patients through a Novel ICT Infrastructure, Journal of Biomedical Informatics, 101: 103342.

[4] S. Schaller, V. Marinova-Schmidt, M. Setzer, et al. (2016) . Usefulness of a tailored eHealth service for informal caregivers and professionals in the dementia treatment and care setting: the eHealthMonitor Dementia Portal. JMIR research protocols, 5(2), p.e47.

[5] A. Papakonstantinou, H. Kondylakis, E. Marakakis, MediExpert: Differential Diagnosis for Educational Purposes, EAI Endorsed Tranactions on e-Learning, 2020.

[6] W. Knaus, D. Wagner, E. Draper, et al. (1991). The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill-hospitalized adults. Chest, 100(6), 1619-1636.

[7] J. Bury, C. Hurt, C. Bateman (2002). LISA: A Clinical Information and Decision Support System for Childhood Acute Lymphoblastic Leukaemia. AMIA.

[8] E. Vardell, M. Moore (2011). Isabel, a clinical decision support system. Medical Reference Services Quarterly, 30(2), 158-166.

[9] J.S Aikins, J.C Kunz, E.H. Shortliffe, R.J. Fallat (2018). PUFF: An expert system for interpretation of pulmonary function data. *Comp. Biom. Res.*, 16(3), 199-208.

[10] R. Boulme, D. Gonzalez, J.C. Schmit (2004). Storing genotypic resistance data and linking to other clinical information. XV International AIDS Conference.

[11] E. Marakakis, K. Vassilakis, N. Papadakis 2010. Meta-rules and uncertain reasoning for diagnosis of epilepsy in childhood. Expert Systems with Applications, 37, 6979-6985.

[12] E. Marakakis, K. Vassilakis, E. Kalivianakis, S. Micheloyiannis S. (2005). An Expert System for Epilepsy with Uncertainty, Journal of Artificial Intelligence and Machine Learning, Special Issue on AI & Specific Applications, 72-78.

[13] M. Stratigi, H. Kondylakis, K. Stefanidis. (2018). FairGRecs: Fair Group Recommendations by Exploiting Personal Health Information, 29th International Conference on Database and Expert Systems Applications (DEXA), pp. 147-155, 2018.

[14] M. Stratigi, H. Kondylakis, K. Stefanidis. (2020). Multidimensional Group Recommendations in the Health Domain, Algorithms 2020, 13(3), 54.

[15] M. Stratigi, H. Kondylakis, K. Stefanidis (2017). Fairness in group recommendations in the health domain. In 2017 IEEE 33rd International Conference on Data Engineering (ICDE) (pp. 1481-1488). IEEE.

[16] A. Martens, J. Bernauer, T. Illmann, A. Seitz (2001). Docs 'n drugs-the virtual polyclinic: an intelligent tutoring system for web-based and case-oriented training in medicine. Proceedings AMIA, pp 433-437.

[17] H. Kondylakis, D. Plexousakis (2012). Ontology evolution: assisting query migration. In International Conference on Conceptual Modeling, pp. 331-344. Springer, Berlin, Heidelberg.

[18] M.Giannoulis, H. Kondylakis, E. Marakakis (2019). Designing and implementing a collaborative health knowledge system, Expert Systems with Applications, 126, pp. 277-294.

[19] M. Giannoulis, H. Kondylakis, E. Marakakis (2018). COSMOS: A web-based, collaborative knowledge system using ontologies and managing uncertainty, 11th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA), pp. 441-448, ACM.

[20] S. McGee, Simplifying likelihood ratios. J GEN INTERN MED 17, 647–650 (2002). https://doi.org/10.1046/j.1525-1497.2002.10750.x

[21] R. Gonzales, Common Symptoms, Chapter 2 in M. A. Papadakis, S.J., McPhee, and M. W. Rabow, (2015), Current Medical Diagnosis and Treatment, Fifty-fourth edition, Mc Graw Hill Education.

[22] R.H. Seller, A.B. Symons (2012). Differential Diagnosis of Common Complaints, 6th Edition, Saunders.

[23] R.R. Baliga, 250 Short Cases in Clinical Medicine, 3rd Edition, Elsevier Saunders, 2002.

[24] A. Stephenson, M. Mueller, J. Grabinar, 100 Cases in General Practice, 1st Edition, Hodder Arnold, 2009.

[25] A. Pappas, G. Troullinou, G. Roussakis, H. Kondylakis, D. Plexousakis (2017). Exploring importance measures for summarizing RDF/S KBs. In European Semantic Web Conference, pp. 387-403.

[26] Š. Čebirić, F. Goasdoué, H. Kondylakis, D. Kotzinos, I. Manolescu, G. Troullinou, M. Zneika, (2019). Summarizing semantic graphs: a survey. The VLDB Journal, 28(3), pp.295-327.

[27] G. Troullinou, H. Kondylakis, E. Daskalaki, D. Plexousakis, D. (2015). RDF Digest: Ontology Exploration using Summaries. In International Semantic Web Conference (Posters & Demos).