# Using analogical complexes to improve human reasoning and decision making in Electronic Health Record Systems

Christian Săcărea, Diana Şotropa and Diana Troancă

Babeş-Bolyai University, Department of Computer Science, Cluj-Napoca, Romania
{csacarea@math,diana.halita@,dianat@cs}.ubbcluj.ro

**Abstract.** A key ability of human reasoning is analogical reasoning. In this context, an important notion is that of analogical proportions that have been formalized and analyzed in the last decade. A bridging to Formal Concept Analysis (FCA) has been brought by introducing analogical complexes, i.e. formal concepts that share a maximal analogical relation enabling by this analogies between (formal) concepts. Electronic Health Record (EHR) systems are nowadays widespread and used in different scenarios. In this paper we consider the problem of improving EHR systems by using analogical complexes in an FCA based setting. Moreover, we present a study case of analogical complexes in a medical field. We analyze analogical proportions in Electronic Health Record Systems and prove that EHRs can be improved with an FCA grounded analogical reasoning component. This component offers methods for knowledge discovery and knowledge acquisition for medical experts based on patterns revealed by analogies. We also show that combining analogical reasoning with FCA brings a new perspective on the analyzed data that can improve the understanding of the subsequent knowledge structures and offering a valuable support for decision making.

**Keywords:** Formal Concept Analysis, Analogical Complexes, Electronic Health Record Systems

## 1 Introduction

Conceptual thinking, logical reasoning, identifying analogies and by thus comparing resulting patterns are key features of human intelligence. Conceptual thinking relies on conscious reflexion, discursive argumentation and human communication, where the role of a thinking, arguing and communicating human being is central in the act of processing knowledge and its inherent structures [19]. The mathematization of conceptual knowledge is done by Formal Concept Analysis (FCA). The basic skills and means of FCA are, by this mathematization, the basic means and skills of Conceptual Knowledge Processing. Logical reasoning is done by *judgements*, understood as asserting propositions, and are formally mathematized by *Conceptual Graphs* [18]. Analogical reasoning has been discussed by J. Sowa in [16], where an Analogy Engine is described. In that paper, Sowa points out that especially in language understanding, identifying analogies and thus analogical processes provide a greater flexibility than more constrained methods. Moreover, analogy is considered to be a prerequisite of logical reasoning.

The idea to support and enhance the medical system has lasted over time. In order to reach these goals, it is important for clinicians to have the possibility to access patient record information anytime and anywhere. The challenge of providing clinicians of any specialty with an integrated view of the complete healthcare history of the patient has so far proved difficult to meet. Due to this major obstacle, our purpose is to effectively deliver healthcare support for clinicians in order to give them a better perspective over a patient healthcare status.

Digitized health information systems are nowadays widespread as a result of constant development of e-health techniques, advances and methodologies. They are comprising many services and systems, relating healthcare with information technology. Among them, Electronic Health Records (EHR) are one of the frames for patient data communication between healthcare professionals.

Electronic Health Records (EHR) describe the concept of a comprehensive collection of a patient's health and healthcare data. They are electronically maintained information about an individual's lifetime health status and healthcare. It has been intensively studied over the last years whether such electronic systems have improved the medical system and what the benefits and the drawbacks are [5, 8, 14, 17].

The development of Electronic Health Records Systems (EHRs) might be influenced by at least two factors: costs and interoperability. Cost has become a critical factor in healthcare. Technical knowledge and cross-border situations are also an important factor in that influence EHRs improvement. Moreover, the ability of a system to interact with one or more other systems has a huge influence over EHRs.

Traditional models of healthcare services are usually inefficient, being associated with expensive interventions. By using EHR systems one could analyze a variety of data about patient's current health status and medical history (medication, symptoms and diagnostics). Furthermore, Electronic Health Record systems may be used in order to reduce human error and thereby to improve patient safety. One may expect that computerized systems could reduce costs, mistakes or unneeded diagnostic tests.

Analyzing medical data is not a new field of research, due to multiple attempts of interpreting data by using statistical methods, neural networks or decision trees [2]. In our previous research, we have proven, by making use of the effectiveness and the graphical representations of conceptual hierarchies, that FCA and its knowledge processing capabilities are good candidates to solve the knowledge discovery, processing and representation task in EHR systems [4, 12]. Similar techniques were used by other researchers, for instance to data mine and interpret patient flows within a healthcare network [6]. Moreover, in our recent research ([13]), we have managed to offer more insights into how patterns can be extracted from medical data and interpreted by means of FCA and Neo4j, i.e., a graph-based database.

In this paper, we expand our focus and prove that EHR can be enhanced with an analogical reasoning Component. For this, we rely on FCA and analogical proportions, as they have been introduced in [9, 10]. Human reasoning is often based on analogical reasoning, but it is also error prone. However, if a computerized system can compute such analogies based on the whole dataset, it could improve the accuracy of the decision process. One of the key notions of analogical reasoning is analogical proportion, which is a statement of the form *a is to b as c is to d*. This statement expresses a similarity between the relation linking $a$ to $b$ and the relation linking $c$ to $d$. Due to different logical and algebraic studies of analogical proportions, reasoning by analogy has been recently improved and formalized. While the use and applications of analogical reasoning are multiple, it was proven, for instance, that analogical reasoning is a powerful tool for classification [1].

On the other hand, FCA is well known for its effective knowledge discovery algorithms and its expressive power, being capable to unify methodologies and find patterns in large datasets. We apply this paradigm in order to obtain in-depth and highly qualitative knowledge representation of medical data. The key notion of FCA is a formal concept, which refers to a maximal subset of elements having a maximal set of properties (attributes) in common.

In the last years, research has suggested FCA and analogical reasoning are two techniques that can be easily combined in the data analysis process, since they both rely on the idea of similarity [9–11]. In this research, we discuss an Analogical Resoning component grounded on the conceptual landscape paradigm of R. Wille [3, 18, 19]. We consider a particular set of medical data to exemplify the FCA based analogies at work. For this, we identify concepts in a weak analogical proportion, i.e., sets of four formal concepts that share a maximal analogical relation. This research proves the effectiveness of FCA together with analogical reasoning as a suitable mechanism through which Electronic Health Record Systems might be improved, offering thus a valuable reasoning and decision making support for medical experts.

The paper is structured as follows. After introducing some preliminaries on EHR systems, FCA and analogical reasoning, we continue with a brief motivation of our medical data analysis. Then, we describe the experiments through which we emphasize how powerful analogical reasoning is by showing what data can be inferred from the results and to what purpose. We end the paper by presenting our conclusion and some future work.

## 2 Preliminaries

### 2.1 Formal Concept Analysis

Conceptual Knowledge Processing is an approach to knowledge management which is based on FCA as its underlying mathematical theory. FCA has constantly developed in the last 30 years, rapidly growing from an incipient lattice theory restructuration approach to a mature scientific field with a broad range of applications. Nowadays, FCA comprises not only important theoretical advances about concept lattices, their properties and related structures (like description logics, and pattern structures), but also algorithmics, applications to knowledge discovery and knowledge representation, and several extensions (fuzzy, temporal, relational FCA, etc).

The basic data type FCA is using is a formal context. Using concept forming operators, a mathematical structure called concept lattice is build. This concept lattice is used as a basis for further communication and analysis. In the following we briefly recall some basic definitions. For more, please consult the standard literature [3].

A *formal context* $\mathbb{K} := (G, M, I)$ consists of two sets $G$ and $M$ and a binary relation $I$ between $G$ and $M$. The elements of $G$ are called *objects* and the elements of $M$ are called *attributes*. The relation $I$ is called the incidence relation of the formal context, and we sometimes write $gIm$ instead of $(g, m) \in I$. If $gIm$ holds, we say that *the object $g$ has the attribute $m$*.

Concept forming operators are defined on the power set of $G$ and $M$, respectively. For $A \subseteq G$, we define $A' := \{m \in M \mid \forall\ g \in A.\ gIm\}$, and for $B \subseteq M$ we have $B' = \{g \in G \mid \forall\ m \in B.\ gIm\}$. These *derivation operators* provide a Galois connection between the power sets of $G$ and $M$. A *formal concept* is a pair $(A, B)$ with $A \subseteq G$ and $B \subseteq M$ with $A' = B$ and $B' = A$. The main theorem of FCA proves that the set of all concepts of a context $\mathbb{K}$ is a complete lattice and every complete lattice occurs as a concept lattice of a suitable chosen formal context.

Usually data are not binary, so objects might have attributes with some values. A *many-valued* context is a tuple $(G, M, W, I)$, where $G, M$ are sets, $I \subseteq G \times M \times W$ is a ternary relation and for all $g \in G$ and $m \in M$ if $(g, m, w) \in I$ and $(g, m, v) \in I$ then $w = v$, i.e., the value of the object $g$ on the attribute $m$ is uniquely determined.

Many-valued contexts can be binarized using a process called conceptual scaling. By this, we can allocate to every many-valued context a conceptual structure consisting of the many-valued context itself, a set of scales and a conceptual hierarchy.

The conceptual structures encoded in medical data can be extracted using FCA and they can be used as a basis for further communication or knowledge acquisition. Using FCA to construct conceptual maps for medical behavior, the obtained lattice structures of these conceptual maps reflect the characteristics of the medical environment.

### 2.2 Proportional Analogies for formal concepts

Through analogical reasoning one may draw plausible conclusions by exploiting parallels between situations. A key pattern which is associated with the idea of analogical reasoning is the notion of analogical proportion (AP), i. e., a statement between two pairs $(A, B)$ and $(C, D)$ of the form $A$ *is to* $B$ *as* $C$ *is to* $D$ where all elements $A, B, C, D$ are in the same category. This relation express that what is common to $A$ and $B$ is also common to $C$ and $D$, and what is different between $A$ and $B$ is also different between $C$ and $D$. There are numerous examples of such statements in our everyday life, i.e. "Paris is to France as Berlin is to Germany", "30 is for 60 what 25 is to 50". Due to the fact that this relation is involving both similarities and dissimilarities between four objects, an analogical proportion is able to model any complex association.

More precisely, according to [9, 10], an analogical proportion (AP ) on a set $X$ is a quaternary relation on $X$, i.e., a subset of $X^4$. An element of this subset, written $x : y :: z : t$, is read as $x$ is to $y$ as $z$ is to $t$, must obey the following two axioms:

1. Symmetry of "as": $x : y :: z : t \Leftrightarrow z : t :: x : y$
2. Exchange of means: $x : y :: z : t \Leftrightarrow x : z :: y : t$

Analogical proportions [9, 10] can be formulated for numbers, sets, in the boolean case, strings, as well as in various algebraic structures, like semigroups or lattices. In the latter case, we say that four elements (or in the case of a concept lattice, formal concepts) $(x, y, z, t)$ of a lattice are in a *Weak Analogical Proportion (WAP)* iff $x \vee t = y \vee z$ and $x \wedge t = y \wedge z$. The symbols $\vee$, respectively $\wedge$, denote the supremum, respectively the infimum as they are defined in Order Theory. Therefore, if we denote by $O_x$ and $A_x$ the extent, respectively the intent of concept $x$, then, by the main theorem of FCA, the previous conditions are equivalent to

$$A_x \cap A_t = A_y \cap A_z \text{ and } O_x \cap O_t = O_y \cap O_z.$$

In some cases, when one of the four elements of the proportion is not known, it can be inferred from the three other elements. The idea of analogy is to establish a parallel between two situations, i.e., what is true in the first situation may also be true in the second one. However, the parallel between two situations that refer to apparently unrelated domains may be especially rich.

Let now $(G, M, I)$ be a formal context, $O_1, O_2, O_3, O_4 \subseteq G$ sets of objects, $A_1, A_2, A_3, A_4 \subseteq M$ sets of attributes. We define the following subcontext, where X denotes that *all* elements of the respective subsets are related by the incidence relation $I$:

| | $A_1$ | $A_2$ | $A_3$ | $A_4$ |
|---|---|---|---|---|
| $O_1$ | | | X | X |
| $O_2$ | | X | | X |
| $O_3$ | X | | X | |
| $O_4$ | X | X | | |

Table 1: Formal context of concepts

When this pattern is maximal, the subcontext is called *analogical complex*. Moreover, every analogical complex defines a WAP and viceversa. Given an analogical complex, let us denote by $x, y, z$ and $t$ the four concepts building a WAP. Analogical complexes contain correspondences between formal concepts as a whole, i.e., elements of the concept lattice. However, if we need to display analogies between objects and attributes, then another approach, called *proportional analogies* is required. For this, Miclet et al. introduce so called *full weak analogical proportions*, which are a more particular case of WAPs, eliminating what they think are trivial WAPs. A full weak analogical proportion (FWAP) [9] is a WAP $(x, y, z, t)$ iff the four concepts are incomparable for $\leq$ and the quadruples $(x, y, z, t)$ and $(x, z, y, t)$ are not WAPs.

A *proportional analogy between concepts* [9] is a relation $\updownarrow$ defined on $\mathfrak{P}(G) \times \mathfrak{P}(M) \times \mathfrak{P}(G) \times \mathfrak{P}(M)$ written as

$$(O_x \setminus O_y) \updownarrow (A_x \setminus A_y) \updownarrow\updownarrow (O_y \setminus O_x) \updownarrow (A_y \setminus A_x).$$

and derived from a full analogical proportion between concepts. However, in this paper we use a "weaker" version of proportional analogies, by taking into account proportional analogies derived from WAPs and not just FWAPS.

A heuristic algorithm to discover such proportions by inspecting a lattice of formal concepts has been proposed in [10]. We have used analogical proportions and analogical complexes in order to compare objects with respect to their attribute values. Our research is following the analogy-based decision approach, through which medical decisions may be inferred from previously formed medical analogies.

## 3 Motivation

Data mining results are typically difficult to interpret, and much effort is necessary for domain experts to turn the results to practical use. In general, users do not care how sophisticated a data mining method is, but they do care how understandable its results are. Therefore, no method is acceptable in practice unless its results are understandable [7].

The motivation of our research is to use analogical reasoning in order to improve the diagnosis process and to lower the risk of human error. Let's consider the following situation in order to have a better understanding of how analogical reasoning can be of use. A doctor of a certain specialization can only diagnose patient with regard to diseases from that particular specialization. In reality, it is often the case that a patient has a disease at the borderline of two or more specializations. What usually happens in that case is that the patient is sent from one doctor to another and he is subjected to multiple tests and investigations. In the end he might even end up with the wrong diagnosis because the two doctors rarely come together to discuss the case an each doctor has limited knowledge about other specializations. How can analogical reasoning help improve this situation?

Having a large database with patients, diseases and symptoms, we can compute proportional analogies, for instance of the form "symptom A is to disease B as symptom C is to disease D". Assuming that symptom A and disease B refer to the specialization of a doctor, it follows that he knows exactly what the correlation is between symptom A and disease B, for example it is the main symptom of that disease, or it is a secondary symptom of that disease or maybe they are not correlated at all. With this knowledge, the doctor can infer new information based on the analogy, namely how symptom C is related to disease D.

Another example would be the following proportional analogy obtained from the analyzed data: *Osteosclerosis* is to *Autophony, Hearing Loss, Rhinorrhea* what *Acute Pharyngitis* is to *Dysphagia, Fever*. A doctor specialized in throat diseases would know that *Dysphagia* and *Fever* are the principal symptoms of *Acute Pharyngitis*. Therefore, he can conclude that a patient having *Autophony, Hearing Loss* and *Rhinorrhea* as symptoms can be diagnosed with *Osteosclerosis* based on analogical reasoning.

In Section 4 we show how to compute weak analogical proportions and their corresponding proportional analogies for a given dataset. Furthermore, we present a few of the obtained results to show how analogical reasoning can be extended on the whole dataset.

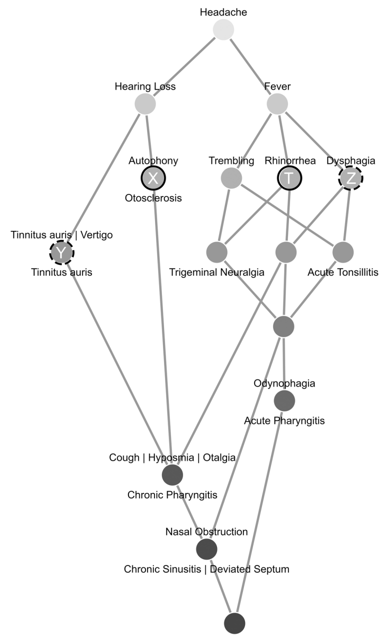
## 4 Experiments

Medical diagnosis is considered one of the most important, and at the same time, complicated task, that needs to be executed accurately and efficiently. The development of an automatic system which may be used as a decision support system which can aid clinicians is therefore of high importance in order to improve accuracy and efficiency.

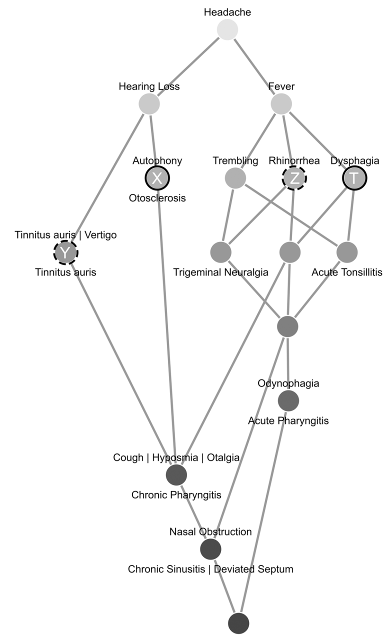
There are multiple situations in which valuable knowledge is hidden in some medical datasets, but it is not directly accesible and quite often this knowledge remains completely inaccessible to medical experts, which need to rely on their expertize, knowledge, experience and sometimes intuition. Of course, this does not necessarily mean that the diagnosis is wrong, but the purpose it may sometimes lead to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients [15]. Data mining techniques have the potential to generate a knowledge-rich environment which can help to significantly improve the quality of clinical decisions. The purpose of such tools is not to replace the doctors' experience and intuition, but to support their decision with facts and knowledge extracted from the data.

We are considering for this analysis records collected from a teaching hospital in Romania. The collected data refer to personal characteristics, symptoms and diagnostics of patients who came to the hospital for an investigation in the Department of Otorhinolaryngology. This department is specialized in the diagnosis and treatment of ear, nose and throat disorders.

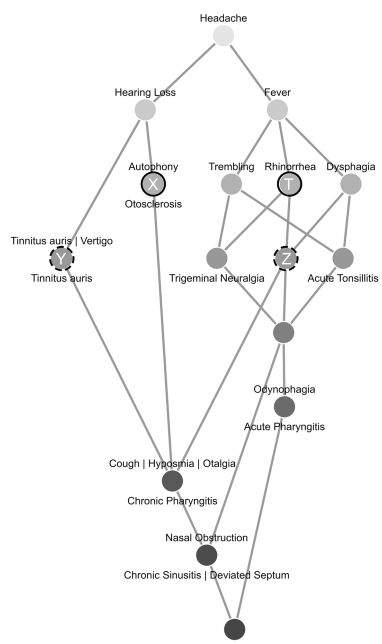
Using these data, we show how new knowledge about medical investigations can be discovered, by following several steps: finding concepts, building knowledge concept lattices, finding relations and analogical proportions between concepts.



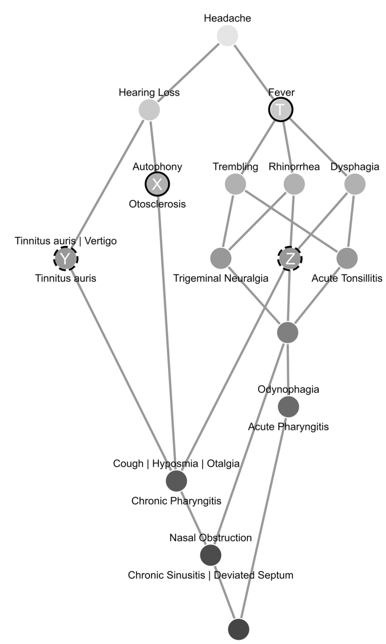

(a) WAP1




(b) WAP2




(c) WAP3




(d) WAP4

Fig. 1: Diagnostics in relation with symptoms with focus on Deviated Septum as principal diagnostic and Chronic Sinusitis as secondary diagnostic

For a better understanding, we detect and represent the analogical proportions on a concept lattice, which has an intuitive and clear representation. Figure 1 presents the conceptual scale for the relation between diagnostics and symptoms which occurred for both male and female patients, but with a particular focus on *Deviated Septum* as principal diagnostic and *Chronic Sinusitis* as secondary diagnostic together with the corresponding symptoms that might appear in patients' data. We are not directly interested in the correlation between the gender of the patient and the diagnostics, but in the cause that led to the diagnostics. For that reason, we look for analogical structures in the concept lattice in order to infer knowledge regarding the correlation between symptoms and diagnostics.

Following the notions introduced in the preliminaries, we look for formal concepts in the lattice that form WAPs, i.e. quadruples of concepts $(x, y, z, t)$, where the two pairs $(x, t)$ and $(y, z)$ share the same infimum and supremum. In this particular case, we identified four quadruples of concepts that fulfill this property. The four WAPs can be observed in Figures 1a, 1b, 1c, respectively 1d, where each WAP is of the form $(x, y, z, t)$, with $x, y, z$ and $t$ being formal concepts (and therefore represented as nodes in the concept lattice). The pairs of concepts sharing the same infimum and supremum are highlighted with different borders for a clear view of the weak analogical proportion.

As explained in the preliminaries, in order to obtain correspondences between diagnostics (objects in the represented formal context) and symptoms (attributes in the represented formal context) we need to compute the corresponding proportional analogies. Herefrom, we can infer similarities between the relations linking different subsets objects to different subsets of attributes. In order to do this we need a clear view of the extents and intents of each concept. Therefore, although the extents and intents can be read directly from the concept lattice, in order to make it easier to follow the computations in the next step, we display them separately in tables. These can be found in Tables 2 - 5 for WAP1 - WAP4, respectively. We observe that WAP2 is only a permutation of WAP1 (obtained by exchanging concepts $t$ and $z$), hence the formal concepts are the same as in Table 2, but with different notations.

When computing the proportional analogies, the definition given by Miclet et al. states that from the WAP formed by the concepts $x, y, z$, and $t$ one can compute the proportional analogy for the pair $(x, y)$. However, given the commutativity of the supremum and infimum operations, we can obviously deduce the same proportional analogies for the following pairs: $(x, z), (y, t)$, and $(z, t)$. As described in the preliminaries, we compute the proportional analogies as set differences of the extents and intents of the detected pairs.

Given these considerations, the proportional analogies for WAP1-WAP4 are displayed in Tables 6-9, respectively. Trivial WAPs, where one of the components is the empty set were eliminated from this analysis.

In this small and relatively easy to read example, we present a few of the weak analogical proportions and their corresponding proportional analogies in order to prove how specific domain knowledge that can be inferred from analog-

| Concept | Extent | Intent |
|---|---|---|
| x | Ostosclerosis | Autophony |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Hearing Loss |
| | Deviated Septum | |
| y | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Hearing Loss |
| | Deviated Septum | Tinnitus |
| | Tinnitus | Vertigo |
| t | Acute Pharyngitis | Fever |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Rhinorrhea |
| | Deviated Septum | |
| | Trigeminal Neuralgia | |
| z | Acute Pharyngitis | Dysphagia |
| | Acute Tonsilitis | Fever |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | |
| | Deviated Septum | |

Table 2: Diagnostic: Deviated Septum - concepts in analogical relation - WAP1

| Concept | Extent | Intent |
|---|---|---|
| x | Ostosclerosis | Autophony |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Hearing Loss |
| | Deviated Septum | |
| y | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Hearing Loss |
| | Deviated Septum | Tinnitus |
| | Tinnitus | Vertigo |
| t | Acute Pharyngitis | Dysphagia |
| | Acute Tonsilitis | Fever |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | |
| | Deviated Septum | |
| z | Acute Pharyngitis | Fever |
| | Chronic Pharyngitis | Headache |
| | Chronic Sinusitis | Rhinorrhea |
| | Deviated Septum | |
| | Trigeminal Neuralgia | |

Table 3: Diagnostic: Deviated Septum - concepts in analogical relation - WAP2

| Concept | Extent | Intent |
|---|---|---|
| x | Ostosclerosis<br>Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum | Autophony<br>Headache<br>Hearing Loss |
| y | Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum<br>Tinnitus | Headache<br>Hearing Loss<br>Tinnitus<br>Vertigo |
| t | Acute Pharyngitis<br>Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum<br>Trigeminal Neuralgia | Fever<br>Headache<br>Rhinorrhea |
| z | Acute Pharyngitis<br>Acute Tonsilitis<br>Chronic Pharingitis<br>Chronic Sinusitis<br>Deviated Septum | Dysphagia<br>Fever<br>Headache<br>Rhinorrhea |

Table 4: Diagnostic: Deviated Septum - concepts in analogical relation - WAP3

| Concept | Extent | Intent |
|---|---|---|
| x | Ostosclerosis<br>Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum | Autophony<br>Headache<br>Hearing Loss |
| y | Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum<br>Tinnitus | Headache<br>Hearing Loss<br>Tinnitus<br>Vertigo |
| t | Acute Pharyngitis<br>Chronic Pharyngitis<br>Chronic Sinusitis<br>Deviated Septum | Dysphagia<br>Fever<br>Headache<br>Rhinorrhea |
| z | Acute Pharyngitis<br>Acute Tonsilitis<br>Chronic Pharingitis<br>Chronic Sinusitis<br>Deviated Septum<br>Trigeminal Neuralgia | Fever<br>Headache |

Table 5: Diagnostic: Deviated Septum - concepts in analogical relation - WAP4

| | Extent1 \ Extent2 | Intent1 \ Intent2 | Extent2 \ Extent1 | Intent2 \ Intent1 |
|---|---|---|---|---|
| **(x,y)** | Ostosclerosis | Autophony | Tinnitus | Tinnitus<br>Vertigo |
| **(x,z)** | Ostosclerosis | Autophony<br>Hearing Loss | Acute Pharyngitis<br>Acute Tonsilitis | Dysphagia<br>Fever |
| **(y,t)** | Tinnitus | Hearing Loss<br>Tinnitus<br>Vertigo | Acute Pharyngitis<br>Trigeminal Neuralgia | Fever<br>Rhinorrhea |

Table 6: Diagnostic: Deviated Septum - proportional analogies - WAP1

| | Extent1 \ Extent2 | Intent1 \ Intent2 | Extent2 \ Extent1 | Intent2 \ Intent1 |
|---|---|---|---|---|
| **(x,z)** | Ostosclerosis | Autophony<br>Hearing Loss | Acute Pharyngitis<br>Trigeminal Neuralgia | Fever<br>Rhinorrhea |
| **(y,t)** | Tinnitus | Hearing Loss<br>Tinnitus<br>Vertigo | Acute Pharyngitis<br>Acute Tonsilitis | Dysphagia<br>Fever |

Table 7: Diagnostic: Deviated Septum - proportional analogies - WAP2

| | Extent1 \ Extent2 | Intent1 \ Intent2 | Extent2 \ Extent1 | Intent2 \ Intent1 |
|---|---|---|---|---|
| **(x,z)** | Ostosclerosis | Autophony<br>Hearing Loss<br>Rhinorrhea | Acute Pharyngitis | Dysphagia<br>Fever |
| **(y,t)** | Tinnitus | Hearing Loss<br>Tinnitus<br>Vertigo | Acute Pharyngitis<br>Trigeminal Neuralgia | Fever<br>Rhinorrhea |

Table 8: Diagnostic: Deviated Septum - proportional analogies - WAP3

| | Extent1 \ Extent2 | Intent1 \ Intent2 | Extent2 \ Extent1 | Intent2 \ Intent1 |
|---|---|---|---|---|
| **(x,z)** | Ostosclerosis | Autophony<br>Hearing Loss<br>Rhinorrhea | Acute Pharyngitis | Dysphagia<br>Fever |
| **(y,t)** | Tinnitus | Hearing Loss<br>Tinnitus<br>Vertigo | Acute Pharyngitis<br>Trigeminal Neuralgia<br>Acute Tonsilitis | Fever |

Table 9: Diagnostic: Deviated Septum - proportional analogies - WAP4

ical reasoning. For instance, by analyzing the pair **(x,z)** from Table 6 and the pair **(x,z)** from Table 7, we can observe that the pair of objects and attributes {{*Ostoclerosis*}, {*Autophony, Hearing Loss*}} is repeated, even if considering different weak analogical proportions. However the second part of the proportional analogy is different, so we can infer different knowledge and enhance the chain of similarities. After analyzing the pair **(y,t)** from Table 6 and the pair **(y,t)** from example Table 7 we can observe that if *Rhinorrhea* is replaced with *Dysphagia* in the symptoms list, there will be replacements in the diagnostics list (*Acute Tonsilitis* will be replaced by *Trigeminal Neuralgia*).

By this small example, we can remark from the data under analysis, on the one hand the real value of the analogical reasoning in this medical setting. For instance, one can deduce that a group of symptoms might describe in a unique way a set of diagnostics. On the other, changing the choice of a single symptom can eventually influence the diagnostic of the patient. For a medical expert, these facts can lead to a better understanding of various patient related data or even to discover eventual inconsistencies in the data under analysis. Moreover, pairs of concepts in weak analogical proportions can be used, combined with the corresponding concept lattice where they are highlighted, for further communication, as an inference support or even as a knowledge discovery method. Facts that stand out should be analyzed on patients' records over a large period of time and, if they persist, it can lead to the formulation of some hypotheses which can than be researched in more details by medical staff for a validation.

Electronic Health Records with an analogical reasoning component enable the discovery of in-depth knowledge which is not always easy to discover. Such a system may address recommendation tasks in order to support clinicians in their medical decisions. Through a small scenario we want to emphasize how a clinician, who has medical experience, can interact with an Electronic Health Record with an analogical reasoning component grounded on FCA:

- Patients comes to the hospital and he / she presents a list of symptoms which are recorded in the system;
- Some of the patient's symptoms are principal symptoms, while others are secondary symptoms;
- Based on a previous data preprocessing phase, conceptual landscapes are built, enabling navigation in a dyadic or triadic setting;
- The analogical reasoning component presents quadruples of concepts in a weak analogical proportion and mines all analogical relations;
- Based on a previously built knowledge base, eventual inconsistencies can be highlighted;
- The clinician can give a diagnostic to the patient and investigates some of the highlighted analogical relations;
- Even more, based on the entire set of analogies, the clinician may dig deeper into relations between symptoms and diagnostics and he / she can specify exactly which diagnostic is the main diagnostic and which are the secondary diagnostics based on the specific list of symptoms.

## 5 Conclusions and future work

Electronic Health Record systems proved to be a significantly aid in health improvement and also a continuous challenge. By means of FCA and analogical reasoning we succeeded in addressing health care system's shortcomings and we have proved that they might have a positive impact on patient welfare.

Through analogical reasoning medical errors can be decreased and due to the existing analogies between patient symptoms and diagnostics, preventing care of a patient may be enhanced.

As future work we propose to develop a recommendation system based on the confidence of an analogy which may be a real help for clinicians. Our perspective is moving on assigning a grading to the recommendation according to the percent of concepts in which the pair $(object, attribute)$ is present out of all concepts which contain the *object* in their extent.

## References


1. Correa, W.F., Prade, H., Richard, G.: Trying to understand how analogical classifiers work. In: Hüllermeier, E., Link, S., Fober, T., Seeger, B. (eds.) Scalable Uncertainty Management - 6th International Conference, SUM 2012, Marburg, Germany, September 17-19, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7520, pp. 582–589. Springer (2012)
2. Delen, D., Walker, G., Kadam, A.: Predicting breast cancer survivability: a comparison of three data mining methods. Artificial Intelligence in Medicine 34(2), 113–127 (2005)
3. Ganter, B., Wille, R.: Formal concept analysis - mathematical foundations. Springer (1999)
4. Haliţă, D., Săcărea, C.: Is FCA suitable to improve electronic health record systems? In: SoftCom 2016. pp. 1–5. IEEE (2016)
5. Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., Taylor, R.: v medical record systems transform health care? potential health benefits, savings, and costs. Health Affairs 24(5), 1103–1117 (2005)
6. Jay, N., Kohler, F., Napoli, A.: Using formal concept analysis for mining and interpreting patient flows within a healthcare network. In: CLA 2006. pp. 263–268 (2006)
7. Li, J., Fu, A., He, H., Chen, J., Jin, H., McAullay, D., Williams, G., Sparks, R., Kelman, C.: Mining risk patterns in medical data. In: SIGKDD 2005. pp. 770–775. ACM (2005)
8. Menachemi, N., Collum, T.: Benefits and drawbacks of electronic health record systems. Risk Management and Healthcare Policy 5, 47–55 (2011)
9. Miclet, L., Barbot, N., Prade, H.: From analogical proportions in lattices to proportional analogies in formal concepts. In: Schaub, T., Friedrich, G., O'Sullivan, B. (eds.) ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014). Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 627–632. IOS Press (2014)
10. Miclet, L., Nicolas, J.: From formal concepts to analogical complexes. In: Proceedings of the Twelfth International Conference on Concept Lattices and Their Applications, Clermont-Ferrand, France, October 13-16, 2015. pp. 159–170 (2015)

11. Miclet, L., Prade, H., Guennec, D.: Looking for analogical proportions in a formal concept analysis setting. In: Proceedings of The Eighth International Conference on Concept Lattices and Their Applications, Nancy, France, October 17-20, 2011. pp. 295–307 (2011)
12. Sacarea, C.: Investigating oncological databases using conceptual landscapes. In: ICCS 2014. pp. 299–304 (2014)
13. Sacarea, C., Sotropa, D., Troanca, D.: Symptoms investigation by means of formal concept analysis for enhancing medical diagnoses. In: Begusic, D., Rozic, N., Radic, J., Saric, M. (eds.) 25th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2017, Split, Croatia, September 21-23, 2017. pp. 1–5. IEEE (2017)
14. Slight, S.P.: Meaningful use of electronic health records: Experiences from the field and future opportunities. JMIR MEDICAL INFORMATICS 3(3), 1–11 (2015)
15. Soni, J., Ansari, U., Sharma, D., Soni, S.: Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications 17(8), 43–48 (2011)
16. Sowa, J., Majumdar, A.: Analogical reasoning. In: Conceptual Structures for Knowledge Creation and Communication, 11th International Conference on Conceptual Structures, ICCS 2003 Dresden, Germany, July 21-25, 2003 Proceedings. pp. 16–36 (2003)
17. Tang, P.C., McDonald, C.J.: Biomedical Informatics: Computer Applications in Health Care and Biomedicine, chap. Electronic Health Record Systems (2006)
18. Wille, R.: Conceptual contents as information - basics for contextual judgment logic. In: Conceptual Structures for Knowledge Creation and Communication, 11th International Conference on Conceptual Structures, ICCS 2003 Dresden, Germany, July 21-25, 2003 Proceedings. pp. 1–15 (2003)
19. Wille, R.: Methods of conceptual knowledge processing. In: Formal Concept Analysis, 4th International Conference, ICFCA 2006, Dresden, Germany, February 13-17, 2006, Proceedings. pp. 1–29 (2006)