A Case Base Reasoning Framework to Author Personalized Health Maintenance Information

Syed Sibte Raza Abidi

Faculty of Computer Science, Dalhousie University, Halifax B3H 1W5, Canada
Email: sraza@cs.dal.ca

Abstract

We present a Personalized Health Information Generation and Delivery System that leverages case based reasoning techniques to dynamically author a Personalized Health Information Package based on an individual's current health profile. The work features a compositional adaptation approach, whereby relevant health information elements from the solution component of multiple similar past cases are carefully selected and systematically combined to yield a new personalized health information package. We have implemented a generic Java-based case based reasoning engine that applies a novel compositional adaptation algorithm to author a HTML-based personalized health information package that can be emailed to users.

1. Introduction

One of the most notable paradigm shifts in healthcare is the preeminence of the wellness maintenance paradigm—a corollary of which is the proliferation of 'generic' healthcare portals over the Internet. However, we note that this medium of healthcare information dissemination has certain limitations, such as a non-specialist individual has to (a) make sense out of generic information targeted towards a general audience; (b) meticulously sift through volumes of healthcare information; and (c) pro-act to find pertinent healthcare information. We argue that the efficacy of web-mediated HI dissemination can be improved by providing *Personalized Healthcare Information*. To achieve maximum impact at the individual-level, HI need to be *personalized* based on an individual's current *Health Profile (HP)* [1]. This is to ensure that the healthcare content disseminated to an individual is specifically focused towards his/her prevailing healthcare needs, akin to the kind of personalized service one enjoys from a visit to a medical practitioner.

Personalization of generic healthcare information, as per an individual's HP, provides an interesting application of CBR [2], in particular the CBR-mediated *Case Adaptation* process [3, 4]. The general idea is to dynamically author a *Personalized Health Information Package (PHIP)* by applying CBR techniques to a cohort of past cases, where each medical-practitioner generated case represents a relationship between the HP component (the situation) and the corresponding PHIP component (the solution/action). In a functional sense, given an individual-specific HP, CBR techniques can be applied to a corpus of past cases—more specifically to the HP component of the cases—to find the best matching past HP(s). The

solution component—i.e. the PHIP—of the retrieved past cases can then be proposed as the candidate PHIP with respect to the given HP. However, in a real-life operational environment, it is rather naive to assume that different individuals may have a similar HP; notwithstanding the possibility that certain HP characteristics may overlap between two (or more) individuals thereby satisfying some coarse-grain similarity criteria, yet there may exist certain HP characteristics that will be idiosyncratic to an individual. For maximum informational accuracy there is an imperative to apply CBR-based *Case Adaptation* techniques to collect 'only' the relevant 'topic-specific' HI documents from pre-composed PHIPs corresponding to the set of matched cases. In this scenario, a new individual-specific PHIP should therefore be compiled by systematically adapting various topic-specific HI documents in line with the current HP.

The rationale of our work is grounded in efforts to provide a technology-enriched solution to individual-focused web-based Health Information (HI) dissemination [3]. In this paper, we present a *Personalised Health Information Generation and Delivery System (PHIGDS)* that leverages advance CBR techniques to dynamically author a PHIP from a corpus of past HP-PHIP type cases. The work features a compositional approach to solution adaptation, designed to address the abovementioned functional requirements. In the realm of PHIGDS, we have developed a generic Java-based *Compositional Adaptation Module* that has the functionality (a) to fine-tune various case adaptation parameters to yield a range of adaptation-driven solutions and (b) to modulate the similarity assessment parameters to practice varying degrees of inter-case similarities. From an operational perspective, PHIGDS can (i) automatically generate an individual's current HP based on information contained in Electronic Medical Record (EMR), (ii) compose an individual-specific PHIP using CBR techniques, and (iii) pro-actively *push* the generated PHIP to the user's email account.

2. Information representation in a CBR framework

Each case comprises two components: (1) the HP component that depicts an individual's health-status defining characteristics, akin to the description of a problem situation; and (2) the PHIP component representing the corresponding solution—the PHIP component is a composite of multiple HI documents (in XML format) that are synthesized to yield a seamless and structured document (in HTML format).

The HP incorporates a substantial volume of up-to-date health-status defining information. We do not treat the HP as a longitudinal list of attributes, rather it has been divided 7 specialized fields; each field corresponds to a particular perspective of an individual's health status (shown in Table 1). The first five fields of the HP—i.e. AD, SI, S, D, A—are regarded as the *core fields* as they describe medically-relevant facts. Whereas the last two fields of the HP—i.e. DD and LD—are deemed as *description fields* as they contain values that describe certain characteristics of an individual, for instance age, gender etc. Note that the PHIP is personalized along the lines of description fields, for example the HI content for an 'adult' and an 'infant' would be different even if they both are suffering from the same acute disease. Hence, the information contained in the description fields modulates the information in the core fields of the HP.

The PHIP is a composite of multiple topic-specific *Healthcare Information Document (HID)*, where each HID contains information about a particular health maintenance topic. In essence, each past case's HP field-value (in the problem-situation component) is associated with a HID. Hence, when a past case's HP field-value is deemed to match with a current HP field-value, the HID associated with it is selected for inclusion in the current PHIP. The logical structure of the PHIP is similar to that of the HP; it comprises of five components, where each PHIP component contains information corresponding to a core field in the current HP. In this

way, each PHIP component contains a list of HIDs that correspond to a set of HP core field values. The amalgamation of all HIDs across the five components realizes the PHIP.

Table 1. Exemplar HP illustrating the 7 HP information groups and their values.

Acute Disease (AD)	Short-Term Illness (SI)	Current Symptoms (S)	Current Drugs (D)	Allerg-ies (A)	Demographic Data (DD)	Lifestyle Data (LD)
Diabetes-Mellitus Rheumatism	Fever Cold	High Temp. Cough Stuffed Chest	Tylenol	Allergic Rhinitis	Age : 45 y Sex : Female Educ: Degree	Fitness: Normal Diet : Healthy Smoking: No

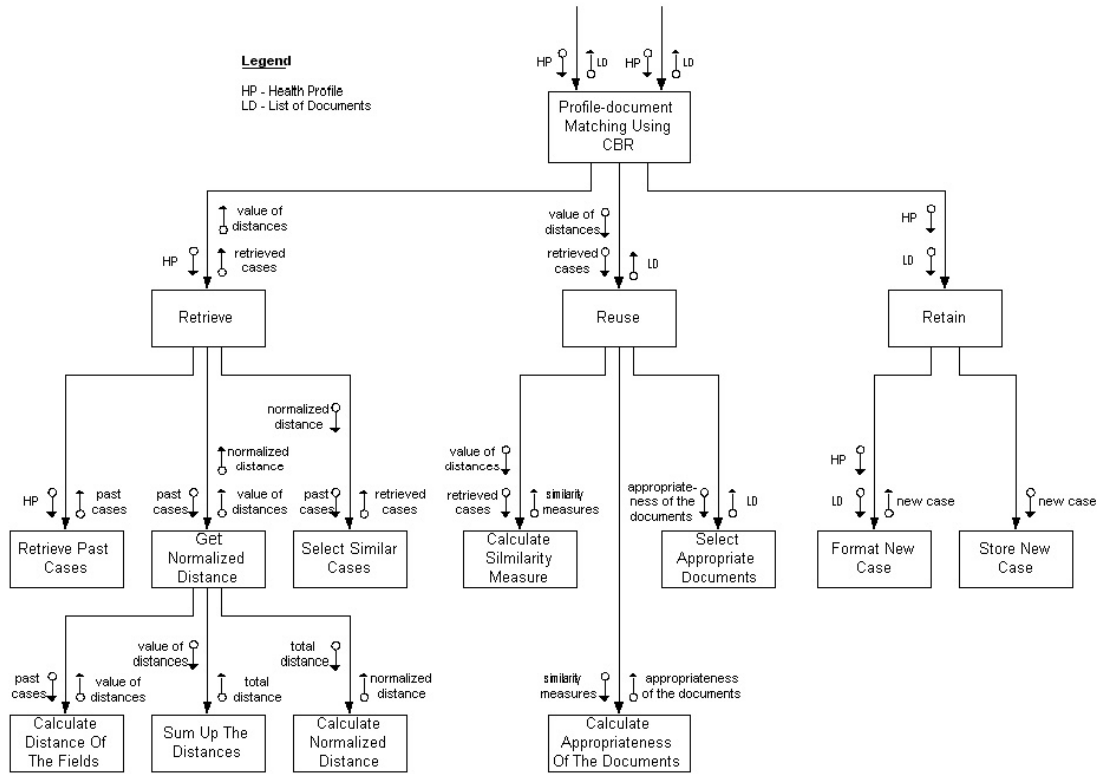# 3. Personalized healthcare information composition

The dynamic PHIP composition involves the following sequence of activities (Figure 1 gives a process flow diagram):

1. Pro-active collection of relevant health status defining information from individualized EMR to derive a current HP. From a CBR perspective, the values of the current HP describe the problem situation (as shown in Tables 1).

2. The current HP forms the basis for performing CBR on a cohort of past cases to derive a new solution—i.e. a dynamically compiled PHIP. In line with CBR methodology, the current HP is compared with the HP component of the past cases. Based on pre-defined similarity assessment criteria a set of closely matched past cases (i.e. both situation and solution components) are retrieved.

3. The solution component of the retrieved cases is systematically adapted to realize an individual-specific PHIP. The solution component—i.e. a list of HIDs—is based on the field-level similarities between the current HP and the past case's HP. Compositional adaptation techniques are used to select the most appropriate HIDs from the available HIDs associated with the retrieved past cases.

4. The list of HIDs are synthesized according to a PHIP template that defines the structural aspects of the PHIP document. Note that the eventual PHIP is a HTML document, whereas the individual HIDs are represented as XML documents.

5. The composed PHIP is finally 'pushed' to the individual via an Internet channel.

3.1. Our case adaptation methodology

We believe that the current HP can be similar to not just one but possibly many HP component of multiple past cases, hence the possibility to combine the solutions of the multiple similar past cases to derive an all-encompassing final solution. In our work of authoring a PHIP we suggest the use of compositional adaptation techniques as they can be extremely useful when the case's solution component is a composite of multiple individual *sub-solutions* [6]. The efficacy of the compositional adaptation approach is inherent in the possibility to select 'relevant' sub-solutions (as opposed to the entire solution) from an ensemble of matching past cases and then synthesizing multiple sub-solutions to derive a more focused, yet elaborate, new solution. In this way, complex problems can be addressed by building upon focused sub-solutions collected from heterogeneous past cases.

Refined Efferent Branch for PHI Delivery System

Legend

HP - Health Profile
LD - List of Documents

HP ⊥ LD HP ⊥ LD

Profile-document
Matching Using
CBR

value of distances
retrieved cases
HP

value of distances
retrieved cases
LD

HP
LD

Retrieve

Reuse

Retain

normalized distance

normalized distance

HP | past cases | past cases | value of distances | past cases | retrieved cases

value of distances | retrieved cases | similarity measures | appropriate-ness of the documents | LD

HP | LD | new case | new case

Retrieve Past Cases | Get Normalized Distance | Select Similar Cases

Calculate Similarity Measure | Select Appropriate Documents

Format New Case | Store New Case

value of distances | total distance

past cases | value of distances | total distance | normalized distance

similarity measures | appropriateness of the documents

Calculate Distance Of The Fields | Sum Up The Distances | Calculate Normalized Distance

Calculate Appropriateness Of The Documents

Figure 1. Process flow diagram showing the CBR-mediated composition of PHIP

Based on the above, we have devised a compositional adaptation approach that is applicable to a specialized class of cases, whereby the case solution is a composite of individual *sub-solutions* each associated with a problem-defining attribute of a case [2]. Specific to our problem description, the case solution—i.e. the PHIP—is also a composite of multiple topic-specific HI documents, each corresponding to a HP-defining attribute.

The premise of our compositional adaptation approach is that for the above class of cases, the solution should not only be derived based on the global similarity between the present and past case(s), rather solution generation should be driven by attribute-level similarity between the present and retrieved past case(s). The rationale for our approach is grounded in the principle that since inter-case similarity is determined at an attribute-level, therefore a fine-grained solution adaptation should be achieved by adapting the attribute-specific sub-solutions based on the attribute's similarity measure derived when comparing the past and present case. As per our approach, for a relatively imperfect match (at the attribute-level) the associated sub-solution will need to be adapted based on the attribute-specific similarity measure between the past and current case [2].

In conclusion, we believe that our approach allows us to more effectively determine the appropriateness of each past case's attribute towards the corresponding present case's attribute. More so, the attribute's appropriateness measure impacts on just the related segment—i.e. the sub-solution component associated with the attribute—of the solution component as opposed to the entire solution component. In this way we are able to generate a more focused solution.

4. Personalized health information generation and delivery system

PHIGDS is implemented as an intelligent info-structure that leverages elements from CBR, Databases and Internet technologies. Figure 2, shows the multi-level architecture of PHIGDS. We briefly explain the functionality of the various modules.
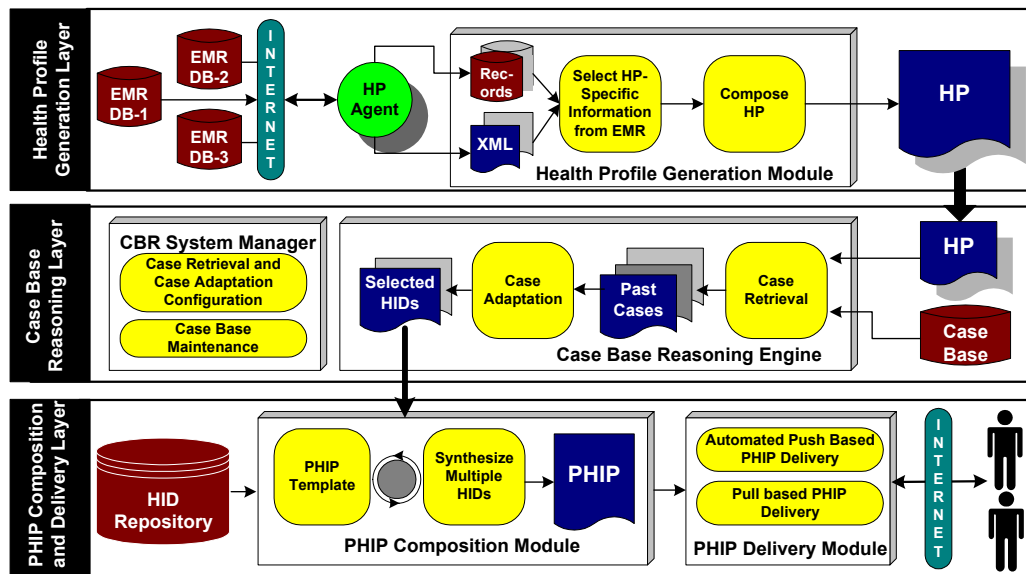
Figure 2. The functional architecture of PHIGDS

The *Health Profile Generation Module* determines the current health profile of an individual by intelligently collating information from Internet-accessible EMR repositories. In practice, PHIGDS routinely scans the EMR to detect any updates to it; the discovery of a recent entry within an EMR initiates the PHIP composition process.

The *Case Base Reasoning Layer* provides the featured CBR techniques employed within PHIGDS. The *CBR System Manager* allows users to configure the various case-retrieval and case-adaptation parameters incorporated in the CBR engine. For instance, field-level weights, case-retrieval and HID appropriateness thresholds.

The *PHIP Composition Module* is responsible to author an individual-specific PHIP from the multiple HIDs selected by the CBR-engine. The multiple XML-based HIDs are synthesized together based on a pre-defined PHIP template—a style-sheet (XSL) that specifies place-holders for imported XML-based text and graphics. The composed PHIP is a seamless XML document, which is subsequently converted to HTML format for Web-mediated delivery to users. Figure 3 illustrates an exemplar PHIP personalized in terms of a given HP.

The *PHIP Delivery Module* features two PHIP delivery modes: (1) *Client-Motivated Mode* that involves the typical *Pulling* of desired information from a web-site and delivery via a web browser; and (2) *System-Motivated Mode* which is the value-added delivery mode whereby up-to-date PHI is pro-actively and periodically *Pushed* to users, over the Internet to their email accounts—i.e. just-in-time PHI automatically delivered at the desktop.

PHIGDS is largely implemented using Java, whereas Javascript has been used to develop the web-based interfaces. The HI content is developed in XML format, with standardized XSL for style and DTD for content specification. The MS SQL server and its JDBC drivers are used to implement the casebase and the HID repository. Finally, we use Jakarta-Tomcat web server for web-mediated access to PHIGDS.

Figure 3. An exemplar PHIP webpage

5. Concluding remarks

In our work we have presented a new approach to compositional adaptation that is applied to the healthcare domain. We conclude that the application of CBR techniques for information personalization—in our case healthcare information personalization—is an interesting alternative to existing personalization techniques. The work presented here is a successful 'proof of concept' and a trial version of PHIGDS is currently deployed for use by a small set of users, with an anticipated increment to the healthcare information content.

References

[1]   S.S.R. Abidi, C.Y. Han, S.R. Abidi, "Patient Empowerment via 'Pushed' Delivery of Personalised Healthcare Educational Content Over the Internet", Proceedings of 10th World Congress on Medical Informatics (MedInfo'2001), September 2 – 5, London. (2001)

[2]   K. Althoff & R. Bergmann, "Case-Based Reasoning for Medical Decision Support Tasks: The INRECA Approach", International Journal of AI in Medicine, Vol. 12 (1998)

[3]   W. Wilke & R. Bergmann, "Techniques and Knowledge Used for Adaptation During Case Based Problem Solving", Lecture Notes in Artificial Intelligence, Vol. 1416. Springer-Verlag, Berlin Heidelberg New York, (1998) pp. 497-505

[4]   N. Arshadi & K. Badie, "A Compositional Approach to Solution Adaptation in Case-based Reasoning and its Application to Tutoring Library", Proceedings of 8th German Workshop on Case-Based Reasoning, Lammerbuckel (2000)