Remote monitoring of chronical pathologies using personalized Markov models

Laurent JEANPIERRE (University Henri Poincaré, Nancy 1)
François CHARPILLET (Directeur de recherche INRIA)
MAIA research team, LORIA

laurent.jeanpierre@loria.fr
Campus Scientifique, BP 239 - F54506, Vandoeuvre-lès-Nancy
Phone: +33(0)383592095; Fax: +33(0)383278319

Abstract. In this paper, an efficient method for diagnosing medical pathologies is presented, which is particularly adapted to long term monitoring of patients. It results from the collaboration between the LORIA (Lorrain Research Laboratory of Computer Science and its Applications), and the ALTIR (Lorrain Association for Renal Failure Treatments). An important aspect of our approach is that the physician can customize each patient's model. This model is expressed in medical terms, simplifying the interaction of physicians with the intelligent system.

The approach will be illustrated with the remote monitoring of patients suffering from kidney disease. In a two years prospective randomized study, 15 patients have been monitored by our system, while 15 others were monitored the classical way. The two groups' statistics has shown that the system was really beneficent to patients' health. This experiment has led to the creation of the DIATELIC enterprise, to promote and develop this system.

Keywords: Probabilistic model based reasoning, human-computer interaction, telemedicine, remote monitoring

1 General considerations

The automated monitoring of a patient is a difficult task, because our knowledge of the dynamics of patient evolution is limited. The remote monitoring is even more difficult, because the monitored patient stays in an uncontrolled environment. Thus, there exists an important noise that spoils data, and alters the patient evolution. For example, a simple stress strongly influences blood pressure.

To cope with so much uncertainty, the choice of probabilistic models seems obvious [1]: their main advantages are their noise tolerance and their ability to model ill-known rules of evolution. More precisely, the class of partially observable Markov models [2] is well adapted to the diagnosis aspect of the problem, because of their observation process. With these models, well-known algorithms are able to uncover the hidden state of the model from the observations. This is an elegant solution to the diagnosis problem, since it only needs a statistical description of the pathologies we are to search for.

Even more important, each patient is unique: even if they almost follow the same generic model [3], there are some important variations hindering the diagnosis process. For example, orthostatic hypotension is a common effect of dehydration; however, this is not true with cardiac insufficient patients. For aging people, arteries tend to lose their elasticity; thus, variations of blood pressure are much different of those of young people.

Finally, the real trouble is that we cannot determine the real condition of a given patient, with very few exceptions. Even the physician may not be sure of his own diagnosis. For example, the only certain symptom of hyper hydration is oedemas presence. But it is already too late: the patient has to go to the hospital. This consideration implies that the model is very hard to define formally. Even measuring the accuracy of the diagnosis is rather difficult, since we cannot know what a good diagnosis is.

All these considerations show clearly that the interaction with physicians has to be a central point in the system. Currently, the physicians' own diagnosis is our only way to define what a meaningful diagnosis is. To have a really working collaboration with them, we chose to ensure a strict medical semantic in each part of the model.

2 Medical context

2.1 General dialysis considerations

The Diatelic project is a real case study of the monitoring of patients suffering from chronic renal insufficiency. In general, this illness is due to kidney diseases which hinder its normal functions [3].

More precisely, our patients are treated through continuous ambulatory peritoneal dialysis. This method allows patients to stay at home while being dialysed daily; they just have to come and see their nephrologist once a month. Between visits, the patient is left alone, even if nurses may come and help.

In this situation, troubles are mainly related to hydration problems: dehydration generally results in comas, while hyperhydration results in increased blood pressure, and may even produce oedemas. High blood pressure can damage the whole organism, and generally worsen kidney diseases rapidly.

The objective of the treatment is then to regulate the patient weight, which is closely related to his hydration level. To achieve this goal, the physicians compute an ideal weight value for each patient. Then, the dialysis strength is modified to reduce the difference between the patient's weight and his ideal weight. The trouble is this ideal weight can also evolve, as the patient grows bigger, or if he loses weight.

In order to enhance these patients' health condition, each patient should be monitored on a daily basis, rather than once a month. Given that each nephrologist often has to monitor more than one hundred patients, they cannot monitor all of them each day. Therefore, we proposed a monitoring architecture [4], in order to detect hydration problems, as well as ideal weight variations. The model we present in this paper is the heart of this system.

2.2 The Diatelic project

Each patient sends his medical data through the Internet daily. These data mainly include the patient weight, his blood pressure, and parameters of his dialysis. Blood pressure is measured when the patient is lit, and when he stands up. The difference between these two measures is named orthostatic tension.

Dialysis parameters include information about each one of the four dialysis bags a patient uses daily. In particular, the system can use bag concentration, its weight, and the time it was injected. After the dialysis, each bag is weighted, and this measure is added to the previous data.

Once a day, each patient uses an Internet connection to transfer those data to a dedicated server. There, these data are stored in a database, and are compared to the patient profile, along with data from previous days. When some anomaly is detected, an alert is sent to the patient, and to the medical team, which will figure out its own diagnosis.

Obviously, all of the gathered medical signals are available to the physicians, along with the computed diagnosis during the last months. With these data, they can see the patient's evolution, and determine if the system was right or if the alert was meaningless. When the patient really is endangered, the treatment may be adapted to overcome this evolution, or the patient might be hospitalized if the standard treatment is insufficient. The objective is to prevent the worsening of patient health, or at least to reduce his recovery duration by detecting troubles before they fully develop.

3 The model

3.1 Markov models

All of the Markov models are based on the Markov property: the behaviour of a system only depends on its current state. These models may be declined as three interesting forms for diagnosing: Markov chains, hidden

Markov models, and partially observable Markov decision processes. Since each model is built on the previous one, I will introduce them from the simpler to the more complex one. Later, I will explain how each of these elements can be applied to our problem.

Markov models are based on a finite set of states, a finite set of actions, and a function of transition. The former describes the evolution of the model state, knowing the current action.

Partially observable Markov models [2] add a new dimension to these models by inserting the observation process. This new stochastic process hides the model state: it simply emits a given observation based on the model state. The challenge is then to infer the hidden state from the observations.

The last model evolution is the partially observable Markov decision processes. They contain another function which is used to reward the system. Hence, an objective is defined, and the system will have to choose actions to reach this goal [5].

Currently, the latest model is not used yet, since we only consider the diagnosis process. However, we could add some rewards if the system was to choose actions to correct the patient state, or simply to get a more accurate diagnosis by asking for complementary tests.

3.2 Medical semantics

The set of states is probably the most important part of the model. In fact, since the objective is to diagnose some pathology, the states of the model must be related to the pathologies we want the system to discover. Classical approaches model the patient state from gathered data, and then try and translate it into an understandable diagnosis.

We chose to introduce the medical semantics directly in the model to enhance collaboration with the medical team. Thus, we have defined our states from a medical point of view. Next, we tried and model each of those states through the gathered data [1]. This approach enhances the physician' s comprehension of the system behaviour, and it also simplifies the system itself since no interpretation is needed.

If we consider the problem of long term monitoring, there is a particular state that plays a crucial role: the “healthy” state. This state is the ideal patient condition, where nothing is wrong. From this point, we can derive other states representing deviations from this healthy condition. Thus, those states will include some pathology, or some conjunction of pathologies.

This model of the patient is relatively rough, since it only knows caricatural situations of static conditions. However, it is generally sufficient for diagnosing simple pathologies. In more complex situations, we may split some state to represent different grades of the health level of the patient. For example, we could split the healthy state in “healthy” and “healthy but stressful”, if this notion was important for the diagnosis. This could even be necessary to ensure that the model satisfies the Markov property. For example, the patient evolution may be biased if he is stressful.

The obvious advantage of this states definition is that each state has a very strong medical meaning. Thus, physicians can interact with the model easily. This helps their cooperation when first designing the model from scratch, and when adapting it to a given patient. Next, it helps them when they want to interpret the computed diagnosis.

The set of actions, combined with the associated transition function models all of the influences the patient can receive. Each action influences the state of the model through a transition matrix. This is the simplest way to implement a probabilistic function indicating how the patient state should evolve from one state to another. From the medical point of view, this transition function is simply the expected evolution of the patient, in response to some treatment (the action). For example, aspirin should lower the overall temperature of the patient. However, in some cases, aspirin is not sufficient; this implies that the state of the model is not precise enough, or simply that, for some unknown reason, the result is not the expected one.

This is why the probabilistic nature of transitions is important: it allows for the use of an approximate knowledge of the patient dynamics. Moreover, physicians are used to work with statistics and those transitions help coping with the noise inherent to the normal life of the patients.

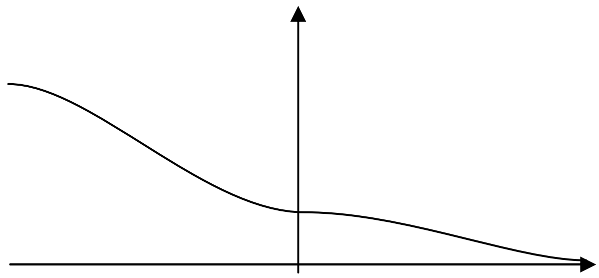
The observation function is the second most important characteristic of the model: it indicates the influence the state of the patient has on the values we can observe. It is the core of the observation process. For example, when a patient suffers from some infection, his ganglia swell. Again, this behaviour is the standard one; there can be exceptions where a patient suffers from infection, and no ganglion swells. It may be related to a factor that has not been modelled into the states; however, it may simply be related to an unknown external influence that prevents the swelling.

For these reasons, the observation function is also probabilistic. It indicates the probability of each possible observation while the patient is in a given state. Different states may have similar probability for some observation. This is why the model state is said hidden: in general, with a given set of observations, no one can determine precisely which the real state of the patient is.

In our model, this function is based on fuzzy sets [6] that allow the model to use continuous observations with relatively few parameters. Thanks to this fuzzy notation, each parameter has an obvious meaning for the medical team. Additionally, this fuzziness allows for a simple but expressive way of describing a given state [7].

Actually, each state is defined separately from the others; additionally, we imposed that data from different sensors were not directly related. For example, weight does not directly influence the blood pressure. Instead, we have an estimation of the hydration level of the patient that influences both the weight and the blood pressure.

This allows for an even simpler collaboration of the nephrologists with the system: each parameter characterizes only a given state, and is only related to a single medical sensor. This influence is represented as a graphical curve indicating how probable a given value of this sensor is, knowing that the patient is in a precise state. For example, blood pressure should be lower than the normal value when the patient is dehydrated; however, in few cases, it may be normal. This will show as the following graph, which is centred on the normal value and expands 1.5 mm Hg each way.



Blood pressure when the patient is dehydrated

Thanks to the strong semantics of states, this kind of graph is really easy to interpret. Moreover, physicians can interact directly with this graph to set the correct influence.

Finally, all of these sensors are aggregated into a single observation probability [8] indicating how representative the physiological signals are, with respect to each state.

4 The diagnosis process

The diagnostic is the art of finding out the reasons that explain observed symptoms. With our model, and more particularly with the state semantics, diagnosing a medical condition is equivalent to finding which the hidden state of the patient is.

To uncover the hidden part of the model from the observations, the traditional algorithm of Viterbi [2] is well adapted. Actually, dynamic programming [9] allows for an efficient use of the Bayes rule for conditional probabilities to obtain a diagnosis process from the declarative model.

Hence, from a given sequence of observations starting from a known state, we are able to compute the optimal sequence of states the patient has visited. More precisely, the Forward algorithm gives the exact probability of any state at each time step, considering all the possible evolutions of the patient.

Once these probabilities have been computed, the physician can interpret them directly, since each state has an obvious medical meaning. This information is far more precise than the mere state sequence: it shows the confidence the system gives to each state. For example, a hesitation between two states is clearly visible, whereas the Viterbi algorithm would have given only the most probable one.

Alerts are generated from these probabilities in two cases: when the most probable state is not the "healthy" one, or if the difference between the two most probable states is negligible. Therefore, alarms are generated if the system is not sure of the state of a patient, or when an anomaly is detected. Next, the alert is displayed on the patient screen to suggest contacting his physician.

At the same time, it is added to the medical team's main screen to draw the physician's attention to this anomaly. From this page, the medical team can see the alerts from all the patients; they can also access all of the gathered data from the past months, and finally adjust the patient's treatment or his profile.

5 The experiment

In association with nephrologists from the ALTIR (Lorrain Association for Renal Failure Treatments), we conducted a prospective randomized experiment during 2 years, with 30 voluntary patients spread across 2 groups. The first one (the Test group) is monitored the classical way, while the second one (the Diatelic group) is monitored with the Diatelic system. Each patient was treated by peritoneal dialysis for one month at home before his integration into the experiment. This way, each patient had time to learn the medical procedure without biasing the experiment results.

Gladly, the two groups were statistically homogenous. At the very beginning of the experiment, all the patients were around 70 years old, with a Charlson index around 5.4. After two years, 12 patients died, and 6 had left the experiment for other reasons. There is no significant difference between the reasons for departing from the two groups.

The first significant difference between the 2 groups is related to the number of visits in excess from the monthly one. Actually, the test group has paid almost 90% more visits per patient than the Diatelic group. This gives an uncertainty factor of 0.66% (ANOVA test). Thus, this diminution is very significant.

The most important difference, from a medical point of view, is that the Diatelic group has a far better controlled blood pressure. In fact, whereas almost all of the patients were slightly in hypertension, patients from the Diatelic group have decreased their blood pressure by a mean 1.1 mm Hg more than the test group ($p<3\%$). In the same time, they slightly decreased their pills consumption ($p<6\%$).

Even more, the average duration of hospital treatment is almost halved: whereas a patient from the Test group stays for 20 (+/- 36) days at the hospital in a year, a patient from the Diatelic group only stays for 11 (+/- 14.5) days at the hospital in a year. Unfortunately, this difference is not statistically significant, principally because of the large standard deviations of both groups.

All these facts amount to an average 14,000 euros annual economy per patient, which largely overcome the price of equipping each patient with a computer. Unfortunately, the costs of continuous ambulatory peritoneal dialysis are mainly related to hospital treatments. Since these are not significant, economies are not significant either.

Considering all the advantages of the system, we decided to patent it [10] and to create an enterprise to continue its development. This year, the experiment will be extended to 300 patients to study the system applicability to a realistic number of patients.

6 Method discussion

We have compared the diagnosis obtained with partially observable Markov models with other techniques. However, it is difficult to quantify exactly the accuracy of a diagnosis. Therefore, this comparison was based on the physicians' remarks on the diagnosis accuracy and on the number of false alerts.

The first method we tested was a standard expert system [11] based on rules we established in collaboration with the ALTIR nephrologists. The model was very hard to tune because of the several thresholds necessary to implement each rule. The worst consequence of this system was that its diagnosis was too sensitive to input variations.

To overcome this sensitivity, we upgraded this system with fuzzy logic [12]. In particular, the sensors' model was exactly the same as the one we used in the Markov model. Each sensor was based on three fuzzy values depending on the signal value being normal, insufficient or excessive. Moreover, each rule was given a confidence factor based on the importance the physicians gave it into their own diagnosis.

However, even this new system was deceptive. In fact, fuzzy logic brought some precision into the observation process, which was a great enhancement. However, it failed into ensuring a temporal stability of the diagnosis. More precisely, the diagnosis of the past days had disproportionate importance: sometimes the diagnosis changed totally from one day to another, sometimes it kept the same diagnosis even if the input data had changed a lot. The trouble is this threshold was very sensitive, and different from one patient to another.

A definite trouble of these two expert systems was that the medical state of the patient was artificially reported from one day to the next one. More than the insertion of several new thresholds to tune this influence, this way of ensuring some temporal stability complicated the rules set beyond what we expected. The result was such that no one was able to read these rules anymore.

Finally, the model parameters were spread all over the rules code, even if we paid particular attention to regroup most of them. The result of this was that it was almost impossible to personalize the patient profile.

During the past year, we compared our results with those obtained with dynamic Bayesian networks [13]. However, even if this model seems to give comparable results in laboratory, no real case use was attempted. Therefore, we cannot say if it is suited to the collaboration with physicians yet.

7 References

1 Jeanpierre, L.: “Apprentissage et adaptation pour la modélisation de processus stochastiques reels”, PhD thesis, University Nancy 1, 2002.

2 Rabiner, L.R.: “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, Proceedings of the IEEE, 1989, 77(2), pp. 257-285.

3 Durand, P-Y., Kessler, M.: “La dialyse péritonéale automatisée, (Massons edn, 1998)

4 Jeanpierre, L., Charpillet, F.: “Hidden Markov Models for Medical Diagnosis”, International Workshop on Enterprise Networking and Computing in Health Care Industry, HEALTHCOM’2002, June 2002, Nancy, France.

5 Cassandra, A.,. Kaelbling, L.P., Littman, M.: “Acting optimally in partially observable stochastic domains”. National Conference on Artificial Intelligence, AAAI 1994, pp 1023-1028.

6 Zadeh, L. A.: “Fuzzy Sets”, Information and Control, 1965, 8, pp 338-353.

7 Steimann, F.: “On the use and usefulness of fuzzy sets in medical”, Artificial Intelligence in Medecine, Vol 21, 2001.

8 Koenig, S., Simmons, R.G.: “Unsupervised learning of probabilistic models for robot navigation”, International Conference on Robotics and Automation, ICRA 1996, pp 2301-2308.

9 Puterman, M.: “Markov Decision Processes: discrete stochastic dynamic programming” (John Wiley & Sons publishers, 1994).

10 Hervy, R., Romary, L., Charpillet, F., Pierrel, J-M, Thomesse, J-P, Petitjean, E., Jeanpierre, L., Durand, P-Y, Chanliau, J.: ‘Système pour le suivi à distance de patients”, Patents in France (FR2804265, 07/27/2001), in Europe (WO0154571, 08/02/2001), and in the U.S.A. (09/539 988, 03/30/2000)

11 http://www.ghg.net/clips/CLIPS.html
Web site of the CLIPS expert system.

12 http://www.iit.nrc.ca/IR_public/fuzzy/fuzzyClips/
Web site of the fuzzy extension of CLIPS.

13 Bellot, D.: “Fusion de données avec des réseaux bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine”. PhD thesis, University Nancy I, 2002