Probabilistic predictive monitoring with CHEERUP

Silvano Mussi

***Abstract*—**The paper presents CHEERUP: a general software environment for building, using and administering application-oriented probabilistic predictive monitoring systems (in the paper called "portals"). Such specific "portals" are used to monitor populations of subjects and get, for single subjects, probabilistic predictions about the occurrence of given undesired/desired events. Probabilistic predictive monitoring is a powerful tool for supporting decisions. It allows to take suitable measures in advance, measures aiming at preventing/favoring the occurrence of the undesired/desired event the application is centered on.

***Keywords*—**Computer applications, Computer engineering, Decision support systems, Predictive monitoring.

I. Introduction

THE possibility of getting early warnings before an undesired event may occur has always been very appealing. Let us think, for example, of prevention of high risk events for health, or serious faults or anomalies of costly and strategic industrial equipments or plants. Similarly, the possibility of getting predictions about the occurrence of a desired event is useful for taking suitable measures in order to favor the event occurrence. Let us think, for example, of passing an exam or reaching a certain athletic performance in the sport field.

The proposal concerns predictive monitoring applied to both preventing undesired events and favoring desired events. The paper presents a general software environment for building and using application-oriented predictive monitoring tools.

The proposal presented in the paper has the ultimate purpose that is in common with many predictive monitoring applications (section V), but, at the same time, it has many aspects that distinguish it from them. In fact CHEERUP is not a predictive monitoring application, it is a general environment for building, using and administering specific applications, i.e. specific application oriented predictive monitoring tools (in the following called portals). CHEERUP can be applied to a great number of heterogeneous domains (Education, Sport, Cultural heritage, Environment, Medicine, Natural Sciences, Social Sciences, Industrial Technology, Economy). In general, it applies to real world domains that can be represented as instances of the following paradigm. There is a population of subjects (human beings, machines, etc.). There is an event E (undesired or desired) that may happen or not to each subject of the population. The occurrence probability of E for a subject may be affected by both the mere aging of the subject and the contexts (i.e. conditions) in which the subject ages. A context has a set of the possible states, as a variable has a set of possible values. The subjects are monitored at constant time intervals by a domain expert. During the monitoring session of a subject the expert enters both the fact "E has occurred/not-occurred" and, for each context, the proper state in which the time has elapsed. In case E = not-occurred CHEERUP simulates, for the subject being monitored, the persistence in the future of a certain set of context states (set defined by the user) and provides a probabilistic prediction about E occurrence in the future, so to help the user take suitable measures in advance. The possibility of simulation supports decision making especially in case of having to choose the more opportune measure given a situation of trade-off.

Moreover let us notice that CHEERUP might also be useful for studying if a given context (or a given combination of contexts) appears to be relevant or not with respect to the probability of E occurrence as time elapses.

CHEERUP facilitates co-operation among work-groups, providing several facilities useful to work in team, in a structured organization.

CHEERUP is a product easy to use. It provides many functions for making its use easy, friendly and proper. It is written in Asp and uses the database management system Mysql.

Even if it is still in a prototype version (for example, the graphic aspect should be improved), it can be effectively applied to real world problems already.

CHEERUP has been conceived and carried out by the author of this paper.

II. CHEERUP Basic Structure

CHEERUP is a general environment which is in turn structured in five target environments (Fig. 1). The first two environments concern the construction and use of specific portals. The remaining three environments concern administration activities.

Manuscript received January 10, 2012: Silvano Mussi is member of AICA (Associazione Italiana per l'Informatica e il Calcolo Automatico) – Piazza Rodolfo Morandi 2, 20121, Milano, Italy. Author e-mail: mussisilvano@tiscali.it.

A. Portals Building Environment

The environment for building portals for specific applications, for short, the *Portals Building environment*, provides a set of functions for building a predictive monitoring portal in a friendly and effective way.

Among the numerous functions the portal builder is provided with there is the possibility of assigning a context state the qualification of *time-sensitive*. This means that the

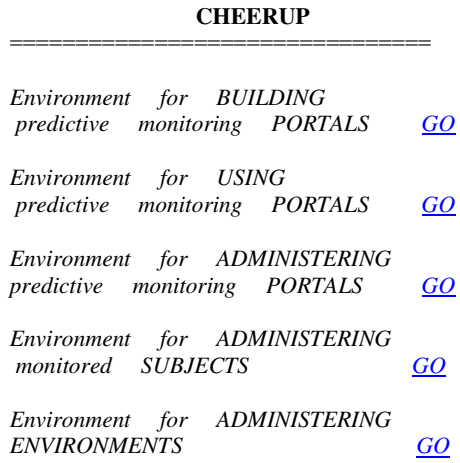

Fig. 1 The Home-Page of CHEERUP with the five top-level environments

portal builder wants that the text of the state explicitly includes how long the subject has elapsed in that context state. The text of the state of a time-sensitive state is automatically extended with the string (called temporal-part): "and such state has lasted for <N time-unit>", where N is an integer number ranging from 0 to the final age considered in the monitoring process.

Another interesting function is represented by the possibility of testing the portal under construction before declaring it definitely finished and ready to be used for real world problems. Such a possibility is very useful since the testing phase might reveal some imperfections of the portal, imperfections that can then be removed by resuming the building phase.

B. Portals Using Environment

The environment for using portals, for short, the *Portals Using environment*, allows the user to use a specific portal built in the Building environment. The Portals Using environment is the one that has to do with the ultimate purpose of CHEERUP: monitoring and predicting. A typical monitoring session of a subject is structured in four basic steps: acquisition of subject data, probabilistic inference about the future occurrence of the undesired/desired event, presentation of probabilistic prediction, session termination and learning.

The environment provides the user of a portal with an interesting possibility. During a monitoring session the portal user can simulate that the subject being examined elapses the future time interval between the time of the present session and the time of the last future session, under a certain combination of context states. More precisely, he/she defines a combination of context states and then asks the Portal Using environment to calculate what would be the future probability of E occurrence if the present combination of context states kept on persisting even in the future. The possibility of such a simulation turns out to be useful, especially in case of having to choose, under trade-off conditions, the best measure to be taken in advance.

Moreover the user is also provided with the possibility of understanding why the probabilistic predictions are what they are. A special explanation function shows the probabilistic reasoning of the Portals Using environment and illustrates where the prediction probability values come from.

C. Administering environments

The environment for administering portals, for short, the *Portals Administering environment*, provides the portal administrator with several utility functions for administering portals that are in the Using environment.

The environment for administering subjects, for short, the *Subjects Administering environment*, provides functions for administering subjects (i.e. the subjects to be monitored).

The environment for administering the environments (i.e. the four environments so far examined), for short, the *Environments Administering environment*, is used by the Super-administrator only. The Super-administrator plays the role of general supervisor of CHEERUP.

III. CHEERUP THEORETIC MODEL

This section presents the mathematical foundation of the Prediction Engine of CHEERUP. It is organized as follows. Subsection 3.1 presents the general theoretic paradigm underlying the Prediction Engine. Subsection 3.2 presents the instantiation of the general theoretic paradigm. Subsection 3.3 illustrates the learning features included in the Prediction Engine. Finally section 3.4 presents the Prediction Engine algorithm.

A. Prediction Engine general paradigm

The basic theoretic paradigm used by CHEERUP is the Dynamic Bayesian network. A Dynamic Bayesian network is basically a Bayesian network [15] in which some links (called "temporal links") represent time elapsing. Many real world domains need to take into account time elapsing. For some variables (i.e. network nodes) the probability distribution on their states is not constant in time. It varies due to the only fact that time elapses. In the real world, time elapses in a continuous way, whereas in a Dynamic Bayesian network it elapses in a discrete way: as a sequence of time-slices. Temporal links allow to represent the effect of time elapsing between two time-slices. In CHEERUP the general model of the Dynamic Bayesian network has been instantiated in the

flowing way. Each monitoring session takes place in a respective time-slice. In each time-slice the event E is represented by a node. The E nodes, present in the respective time-slices, are connected by temporal links. Given two time-slices: t1 and t2, (t2 > t1) and the event E, it can be stated that the value of P(E=occurred) in t2 might be different from the value of P(E=occurred) in t1 for the only fact that an interval time equal to t2 – t1 has elapsed. The conditions in which a subject passes the time interval between t1 and t2, are represented by selecting the appropriate context states in the time slice t2. If in the real world the event E occurs in the time interval between t1 and t2, in the model the event E occurs in the time slice t2, and its occurrence is represented by selecting the state "E=occurred" in the time slice t2. Moreover in the time slice t2 are also selected the context states representing the conditions in which the subject has spent the first part of the time interval before the E occurrence. Obviously the model involves a reality approximation, approximation that is as smaller as temporally nearer sessions are.

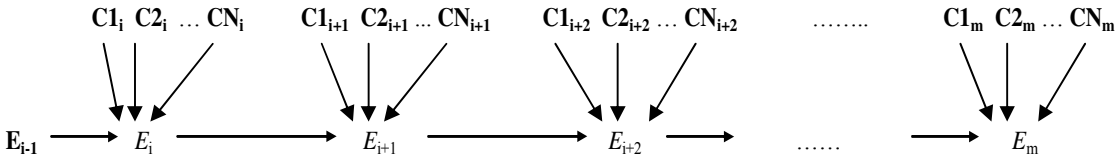

Fig. 2. The structure of the Bayesian network used to produce predictions. Session i is the first future session (i ≥ 2) and session m is the last future session. The arrows connecting E nodes represent temporal links. The nodes in bold are instantiated, the probability values of the E nodes in italic (not in bold) are to be calculated

B. Prediction Engine network

In CHEERUP the mathematical model of prediction consists in a Dynamic Bayesian network that is dynamically built and executed when the portal user asks the Portals Using environment for predictions, that is, in other words, when the user asks: what would be the future probability of E occurrence if the present combination of context states kept on persisting in the future too? Let us examine the basic structure of such a network. The contexts defined by the portal builder: C1, C2, …, CN, and the event E are the nodes of the network. Let us use the symbols E_i and $C1_i$, $C2_i$, …, CN_i to respectively denote the event E and the contexts C1, C2 ,…, CN related to a session i. Let us use the symbol "→" to represent a causal link, so that "A → B" means "A causes B". Let session i and session m be the first future session and the last future session respectively, obviously m ≥ i and i ≥ 2 (that is the first session cannot obviously be the first future session). The Bayesian network used to produce predictions for each session k between i and m (i.e. i ≤ k ≤ m) is the one shown in Fig. 2. For short, let us use "y" and "n" to denote "occurred" and "not-occurred" respectively. Let us notice that E_{i-1}=n (in fact if it were E_{i-1}=y the Portals Using environment would communicate to the user that there is no future session to simulate). Let us assume that for a session k >1, if E_{k-1}=y, then E_k=y independently of the combination of context states in session k. In fact, if in a certain time-slice, E is assigned the state "occurred", in a subsequent time-slice E will obviously keep the same state. However in practice the event E=occurred causes the end of the monitoring process, that is if for a given subject E_i=occurred, then the subsequent sessions i+1, i+2, … do not exist any more for that subject.

C. Prediction Engine learning

Before entering the prediction engine core let us premise that whenever a monitoring session ends, a learning process takes place. Let us examine it more precisely. For short let us represent an instantiation of contexts (i.e. assigning each context a proper state) related to a session k by simply writing $C1_k$, …, CN_k (instead of $C1_k$=st1, …, CN_k=stn, where st1 and stn are elements of the sets of states of C1 and CN respectively). Let k be the current session and let $C1_k$,…, CN_k be the contexts instantiation for session k. At the end of session k the learning process consists in updating the value of

$$P(E_k = y \mid C1_k,...,CN_k) \quad (1)$$

in case of k=1 (i.e. in case of the first session of the monitoring process), or

$$P(E_k = y \mid E_{k-1} = n, C1_k,...,CN_k) \quad (2)$$

in case of k >1.

By adopting the Frequency Probability definition, it can be stated that, if k = 1,

$$P(E_1 = y \mid C1_1,...,CN_1) = \frac{N_{1y}}{N_{1tot}}$$

where N_{1y} is the number of cases that in session 1 have been found with E_1=y | $C1_1$,…, CN_1 whereas N_{1tot} is the total number of cases so far examined in session 1 given the instantiation $C1_1$,…, CN_1. Obviously such a ratio has to be intended as an empirical probability value approximating the theoretical probability value, approximation that is as smaller as greater N_{1tot} is.

Similarly, if k > 1,

$$P(E_k = y \mid E_{k-1} = n, C1_k, ..., CN_k) = \frac{N_{ky}}{N_{ktot}}$$

where N_{ky} is the number of cases that in session k have been found with E_k=y | E_{k-1}=n, $C1_k$,…, CN_k whereas N_{ktot} is the total number of cases so far examined in session k given the instantiation $C1_k$,…, CN_k. In conclusion, learning is accomplished by bringing up to date the numbers: N_{ky} , N_{ktot} and consequently the quotient N_{ky} / N_{ktot} . For short, let us denote with L_k the value learned by means of the (1) if k=1, or the (2) if k>1.

D. Prediction Engine algorithm

After the presentation of the network and the learning process of the Prediction Engine we are now ready to examine the algorithm of the Prediction Engine so to understand more deeply what predictions consist in. If session i is the first future session (necessarily i ≥ 2) and session m is the last future session, CHEERUP probabilistic predictions consist in calculating, for each future session k, where i ≤ k ≤ m, the value of

$$P(E_k = y \mid E_{i-1} = n, C1_i, ...CN_i, ..., C1_k, ..., CN_k)$$

For k=i such a value is L_i , the value learned by the environment, according to the (2). Let us now face the problem concerning the case of i < k ≤ m. For short let us use the symbol A to denote the sequence:

$$E_{i-1} = n, C1_i, ..., CN_i, ..., C1_k, ..., CN_k$$

It can be stated that

$$P(E_k = y \mid A) =$$
$$P(E_k = y \mid E_{k-1} = n, A) \cdot P(E_{k-1} = n \mid A) + \quad (3)$$
$$P(E_k = y \mid E_{k-1} = y, A) \cdot P(E_{k-1} = y \mid A)$$

In fact

1) by applying the product rule we have:

$$P(E_k = y \mid E_{k-1} = n, A) \cdot P(E_{k-1} = n \mid A) =$$
$$P(E_k = y, E_{k-1} = n \mid A)$$

and similarly

$$P(E_k = y \mid E_{k-1} = y, A) \cdot P(E_{k-1} = y \mid A) =$$
$$P(E_k = y, E_{k-1} = y \mid A)$$

2) since the two joint events (E_k=y, E_{k-1}=n) and (E_k=y, E_{k-1}=y) are mutually exclusive, on the basis of the addition axiom we have:

$$P(E_k = y, E_{k-1} = n \mid A) + P(E_k = y, E_{k-1} = y \mid A) =$$
$$P((E_k = y, E_{k-1} = n \mid A) OR (E_k = y, E_{k-1} = y \mid A))$$

3) since the set of states {E_{k-1}=n , E_{k-1}=y} is exhaustive, we have:

$$P((E_k = y, E_{k-1} = n \mid A) OR (E_k = y, E_{k-1} = y \mid A)) =$$
$$P(E_k = y \mid A)$$

On the basis of these considerations let us rewrite the (3) as follows (for short the sequence $C1_i$, …, CN_i, …, $C1_k$, …, CN_k is represented by $C1_i$, …, CN_k):

$$P(E_k = y \mid E_{i-1} = n, C1_i, ..., CN_k) =$$
$$P(E_k = y \mid E_{k-1} = n, E_{i-1} = n, C1_i, ..., CN_k) \cdot \quad (4)$$
$$P(E_{k-1} = n \mid E_{i-1} = n, C1_i, ..., CN_k) + \quad (5)$$
$$P(E_k = y \mid E_{k-1} = y, E_{i-1} = n, C1_i, ..., CN_k) \cdot \quad (6)$$
$$P(E_{k-1} = y \mid E_{i-1} = n, C1_i, ..., CN_k) \quad (7)$$

Let us consider the (4). Every causal path connecting the nodes E_{i-1} , $C1_i$,…, CN_i ,…, $C1_{k-1}$,…, CN_{k-1} to the node E_k is a serial structure in which E_{k-1} is the last but one node. Since E_{k-1} is instantiated to a state (i.e. the state n), each of its antecedents (that is the nodes E_{i-1} , $C1_i$,…, CN_{k-1}) does not affect E_k so they can be neglected (see Bayesian network theory) and therefore the (4) is equivalent to the (2). The ultimate consequence is that the value of the (4) is known: it has been learned by the environment, it is L_k. The value of the (5) is complementary to the value of the (7). The value of the (6) is 1 (as above pointed out). Finally let us consider the (7). The nodes E_{k-1} , $C1_k$,…, CN_k are all direct causes of the node E_k (there is a causal structure converging to E_k). Since E_k is not instantiated to any of its states, its causes are all independent (see Bayesian network theory). Therefore the nodes $C1_k$,…, CN_k does not affect E_{k-1} , and as a consequence they can be neglected. The ultimate consequence is that the (7) is equivalent to

$$P(E_{k-1} = y \mid E_{i-1} = n, C1_i, ..., CN_{k-1}) \quad (8)$$

But the value of the (8) is the prediction value calculated for session k-1. So, in general, it can be stated that:

$$P(E_k = y \mid E_{i-1} = n, C1_i, ..., CN_k) =$$
$$L_k \cdot (1 - X_{k-1}) + X_{k-1} \quad (9)$$

where X_{k-1} stands for the prediction value calculated for session k-1. In conclusion, the value of

$$P(E_k = y \mid E_{i-1} = n, C1_i, ...CN_i, ..., C1_k, ..., CN_k)$$

produced by the Prediction Engine algorithm is given by

L_i if k=i,

$L_k \cdot (1 - X_{k-1}) + X_{k-1}$ if i < k ≤ m.

Let us notice that if k=i+1 the (9) is instantiated as follows

$$P(E_{i+1} = y \mid E_{i-1} = n, C1_i, ..., CN_{i+1}) =$$
$$L_{i+1} \cdot (1 - L_i) + L_i \quad (10)$$

So far we have presented the mathematical model currently implemented in CHEERUP. Let us notice though that an equivalent Prediction Engine algorithm might be defined by following an alternative approach: the Chain Rule based approach. Given a Bayesian network, the probability values of its nodes can be calculated by the global joint probability table of the network. Such a table is built by applying the Chain Rule (see Bayesian network theory). Let us consider the Bayesian network of figure 1. For the sake of simplicity let us suppose for the moment that m=i+1. For short, let us represent the sequence $C1_i$,…, CN_i ,…, $C1_{i+1}$,…, CN_{i+1} by simply writing $C1_i$,…, CN_{i+1}. The value of

$$P(E_{i+1} = y \mid E_{i-1} = n, C1_i, ..., CN_{i+1})$$

is calculated from the global joint probability table by simply summing the values of all the rows containing both $E_{i+1} = y$ and $E_{i-1} = n, C1_i, ..., CN_{i+1}$.

As a consequence we have:

$$P(E_{i+1} = y \mid E_{i-1} = n, C1_i, ..., CN_{i+1}) =$$
$$P(E_{i+1} = y, E_i = n, E_{i-1} = n, C1_i, ..., CN_{i+1}) +$$
$$P(E_{i+1} = y, E_i = y, E_{i-1} = n, C1_i, ..., CN_{i+1})$$

By applying the Chain Rule we have that:

$$P(E_{i+1} = y, E_i = n, E_{i-1} = n, C1_i, ..., CN_{i+1}) =$$
$$P(E_{i+1} = y \mid E_i = n, C1_{i+1}, ..., CN_{i+1}) \cdot$$
$$P(E_i = n \mid E_{i-1} = n, C1_i, ..., CN_i) \cdot$$
$$P(E_{i-1} = n) \cdot P(C1_i) \cdot ... \cdot P(CN_{i+1})$$

and, similarly,

$$P(E_{i+1} = y, E_i = y, E_{i-1} = n, C1_i, ..., CN_{i+1}) =$$
$$P(E_{i+1} = y \mid E_i = y, C1_{i+1}, ..., CN_{i+1}) \cdot$$
$$P(E_i = y \mid E_{i-1} = n, C1_i, ..., CN_i) \cdot$$
$$P(E_{i-1} = n) \cdot P(C1_i) \cdot ... \cdot P(CN_{i+1})$$

Let us notice that if E_{k-1}=y, then E_k=y independently of what the instantiation $C1_k$,…, CN_k is. Let us notice that $P(E_{i-1}=n)=1$. Finally let us remember that the symbols $C1_i$,…, CN_{i+1} denote context instantiations (see Fig. 1) and as a consequence: $P(C1_i)=1$,…, $P(CN_{i+1})=1$. By taking into account these considerations it can be stated that

$$P(E_{i+1} = y \mid E_{i-1} = n, C1_i, ..., CN_{i+1}) =$$
$$L_{i+1} \cdot (1 - L_i) + L_i$$

But this is just the above (10), i.e. starting from another approach we have got to the (10).

Let us now generalize for any k, where $i \le k \le m$. We have:

$$P(E_k = y \mid E_{i-1} = n, C1_i, ...CN_i, ..., C1_k, ..., CN_k) =$$
$$L_k \cdot (1 - L_{k-1}) \cdot (1 - L_{k-2}) \cdot ... \cdot (1 - L_{i+1}) \cdot (1 - L_i) +$$
$$L_{k-1} \cdot (1 - L_{k-2}) \cdot ... \cdot (1 - L_{i+1}) \cdot (1 - L_i) +$$
$$...$$
$$L_{i+1} \cdot (1 - L_i) +$$
$$L_i$$

IV. A SIMULATED-CASE STUDY

In order to make the presentation easier to understand, let us make abstract concepts concrete by referring to a simulated example chosen inside a specific domain: the medicine domain. However the reader should not intend that CHEERUP is a proposal suitable to medicine only, the medicine domain is considered just as an example. The example is not developed in a scientific rigorous manner as a physician would do (the terminology too may be imperfect). It appears incomplete, very simplified and/or naive, especially if the reader is a physician. The purpose of such an example is to make even a non physician reader able to get a clear comprehension of how CHEERUP works.

The medicine domain is a typical case in which there are several undesired events whose occurrence is favoured if the subject passes long time in some contexts that are commonly called risk factors. For example, let us consider the event *First Cardiac Infarct*. Among the set of the related risk factors we might identify: *obesity*, *hypertension*, *abnormal cholesterol levels*, *smoke*, etc. Let us suppose that a medical monitoring enterprise to prevent the occurrence of the first infarct has been put into practice for a population of mail subjects starting from a certain age. During a monitoring session the physician takes note of the presence/absence of the considered risk factors. Let us notice though that for some risk factors it might not be enough to know that at session time they result to be present. In fact it might also be necessary to know how long the subject has passed in presence of those risk factors. For example, the longer the subject has been smoking the higher the contribution that smoke gives to rise the occurrence probability of the event *First Cardiac Infarct* is. After this premise let us start to build the *First Cardiac Infarct* portal (in CHEERUP, portals have the same names as the related undesired/desired events) and define the simulated statistical initial-data. These statistical initial-data do not come from a real monitoring process of a population of real subjects. They though allow to simulate a real case, making this way CHEERUP work in a context similar to a real world context. The reader is therefore able to get a deeper and concrete idea of what the CHEERUP prediction facility consists in.

A. *Definition of the simulated application*

The home-page of the Portal Building environment consists in a set of functions that allow the builder to create and edit the various components of the portal. By using these functions it has been built an application with the following simulation configuration data:

1) Event: First Cardiac Infarct (with states: “occurred”, “not-occurred”)
2) Population to be monitored: mail persons
3) Time period of the whole monitoring process: from the age of 60 years (first monitoring session) to the age of 80 years (last monitoring session)
4) Time interval between two consecutive monitoring sessions: 2 years
5) Minimum number of cases needed to draw predictions (threshold value): 100
6) Partition of the probability interval [0,1] (alarm levels): very-low (from 0 to 0.15), low (from 0.15 to 0.30), low-middle (from 0.30 to 0.45), middle (form 0.45 to 0.60), middle-high (from 0.60 to 0.75), high (from 0.75 to 0.90), very-high (from 0.90 to 1)
7) Contexts (i.e. risk factors):
   - Obesity
   - Hypertension
   - Cholesterol (i.e. high levels of cholesterol)
   - Smoke (i.e. cigarette smoke)
8) Context states:

- Obesity, 2 states: s1 = "no" (i.e. absent), s2 = "yes" (i.e. present)
- Hypertension, 2 states: s1 = "no", s2 = "yes"
- Cholesterol, 2 states: s1 = "ok", s2 = "not-ok"

9) States of the 9 states of the context Smoke (TI represents the length of the Time Interval in which the subject had been smoking):
   - s1 = "no and never in the past"
   - s2 (time-sensitive) = "no but yes in the past (for TI > 30 years)"
   - s3 (time-sensitive) = "no but yes in the past (for 25 < TI ≤ 30 years)"
   - s4 (time-sensitive) = "no but yes in the past (for 20 < TI ≤ 25 years)"
   - s5 (time-sensitive) = "no but yes in the past (for 15 < TI ≤ 20 years)"
   - s6 (time-sensitive) = "no but yes in the past (for 10 < TI ≤ 15 years)"
   - s7 (time-sensitive) = "no but yes in the past (for 5 < TI ≤ 10 years)"
   - s8 (time-sensitive) = "no but yes in the past (for TI ≤ 5 years)"
   - s9 (time-sensitive) = "yes"

B. Definition of the simulated statistical data

In order to run the application above defined, we need to simulate that at each session, for each possible combination of context states there is a minimum number of cases necessary for drawing probabilistic inferences. In other words, in order to draw inferences that, although in a coarse manner, simulate reality, we need to define, for each age and for each possible combination of context states, a proper value representing the *initial* percentage of subjects that have had first cardiac infarct.

Let us define the following simulated statistical data for the age 62 (i.e. the age of the second monitoring session):

1) Let us consider a population of 100 subjects
2) Let N_{62} be the number of subjects that at the age of 62 years (i.e. at the second session) have First Cardiac Infarct= "occurred". Let us consider the better context-states combination: Obesity= "no", Hypertension= "no", Cholesterol= "ok", Smoke= "no and never in the past". Let us establish: (N_{62} | Obesity= "no", Hypertension= "no", Cholesterol= "ok", Smoke= "no and never in the past") = 1.
3) Let Smoke be in state "no and never in the past". Taking into account that the contexts above defined are risk factors that favour the occurrence of first cardiac infarct let us define the following rule:
   - if Obesity= "yes", then add 2 else add 0
   - if Hypertension= "yes" then add 2 else add 0
   - if Cholesterol= "not-ok", then add 2 else add 0
4) For example, (N_{62} | Obesity= "yes", Hypertension= "no", Cholesterol= "no", Smoke= "no and never in the past") = 3; (N_{62} | Obesity= "yes", Hypertension= "no", Cholesterol= "not-ok", Smoke= "no and never in the past") = 5, etc.
5) Since the 8 Smoke states: s2, s3, s4, s5, s6, s7, s8, s9, are time-sensitive, they are automatically added with the extension *and such state has lasted for* T *years*, where T is an element of the set of integer values {0, 1, …, 80}.
6) Let us consider Smoke= "yes". Let us notice that the longer a subject has been smoking the higher the occurrence probability of the first cardiac infarct is. For the sake of simplicity let us consider 7 temporal intervals and let us define the following rules:
   - if T ≤ 5 years, then add 2
   - if 5 < T ≤ 10 years, then add 3
   - if 10 < T ≤ 15, then add 4
   - if 15 < T ≤ 20, then add 5
   - if 20 < T ≤ 25, then add 6
   - if 25 < T ≤ 30, then add 7
   - if T > 30, then add 8
7) For example, (N_{62} | Obesity= "no", Hypertension= "no", Cholesterol= "ok", Smoke= "yes" *and such state has lasted for* 8 *years*) = 4; (N_{62} | Obesity= "yes", Hypertension= "no", Cholesterol= "not-ok", Smoke= "yes" *and such state has lasted for* 22 *years*) = 11;
8) Let us consider the 7 Smoke-states: "no but yes in the past (for …TI…)" accompanied with the temporal part *and such state has lasted for* T *years*. The meaning of T in the temporal part, is that the subject, after having smoked in the past for an time interval equal to TI, stopped smoking T years ago. Let us notice that the occurrence probability of first cardiac infarct decreases when the subject stops smoking: after 1 year it decreases to 50%, after 15 years the negative effects (regarding cardiac infarct) of smoking are no longer present. For the sake of simplicity, let us consider only 3 temporal intervals and let us define the following rules:
   - if T ≤ 2 years, then: if s2 then add 7, if s3 then add 6, if s4 then add 5, if s5 then add 4, if s6 then add 3, if s7 then add 2, if s8 then add 1
   - if 2 < T ≤ 10 years then: if s2 then add 2, if s3 then add 2, if s4 then add 2, if s5 then add 1, if s6 then add 1, if s7 then add 1, if s8 then add 0

- if T > 10 years, then add 0 (that is: null negative effect)

9) For example, (N_{62} | Obesity= “yes”, Hypertension= “no”, Cholesterol= “not-ok”, Smoke= “no but yes in the past (for 25 < TI ≤ 30 years)” *and such state has lasted for* 3 *years*) = 7

Let us define the general rule for simulated statistical data for the other ages of the whole monitoring process.

1) Given that the monitoring process starts at the age of 60 years (first session) and ends at the age of 80 years (last session), and monitoring sessions occur every 2 years, we have to do with 11 sessions: subject age ranges from 60 to 80 with step= 2 (e.g. 60, 62, 64, 66, etc.). As for the second session (corresponding to the age of 62 years) the simulated statistics rules have been defined above. Let us now define the general rule for any other session. Taking into account that aging is itself a condition that can favour first cardiac infarct let us define the following rule (where csa stands for “current session age” and sc stands for “state combination”):
   - (N_{csa} | sc) = (N_{62} | sc) + (csa – 62) / 2
2) For example, (N_{66} | Obesity= “yes”, Hypertension= “no”, Cholesterol= “not-ok”, Smoke= “yes” *and such state has lasted for* 22 *years*) = (N_{62} | Obesity= “yes”, Hypertension= “no”, Cholesterol= “not-ok”, Smoke= “yes” *and such state has lasted for* 22 *years*) + (66 – 62) / 2 = 11 + 2 = 13

C. *Execution of the simulated application*

After the portal has been built and the *initial* statistical data (the ones produced by the rules defined in subsection B) have been put into the database, the portal has been used and the prediction engine has been activated with different simulated situations.

1. *Case 1*

Let us suppose we are at the first session (subject age = 60 years) and we have to do with a subject having the best combination: Obesity= ”no”, Hypertension= ”no”, Cholesterol= “ok”, Smoke= “no and never in the past”. Starting from this situation we ask the prediction engine to calculate the
occurrence probability of the first cardiac infarct if this context states combination keeps constant even in the future (i.e. for all the future ages). The prediction engine has calculated predictions by using the *initial* statistical values. It has produced the numeric outcome showed in Table I.

The prediction engine has also produced the qualitative view represented by the histogram showed in Fig. 3.

2. *Case 2*

Let us contrast Case 1 and let us suppose that at the first session the subject has a combination very bad for his/her health: Obesity= ”yes”, Hypertension= ”yes”, Cholesterol= “not-ok”, Smoke= “yes” *and such state has lasted for* 40 *years*.

Starting from this situation we ask the prediction engine to calculate predictions in the hypothesis that this bad situation keeps constant even in the future. The calculus has used the *initial* statistical values and has produced the numeric outcome showed in Table II.

Table I Predictions table related to Case 1

AGE years	PROB	LEVEL
62	0,01	very-low
64	0,0298	very-low
66	0,0589	very-low
68	0,0965	very-low
70	0,1417	very-low
72	0,1932	low
74	0,2497	low
76	0,3097	low-middle
78	0,3718	low-middle
80	0,4346	low-middle

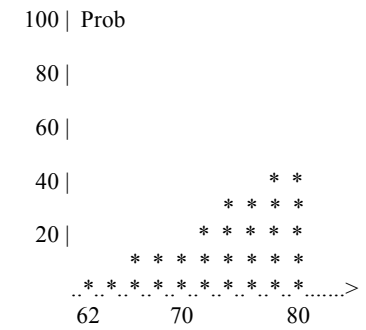

Fig. 3 Predictions histogram related to Case 1

Table II Predictions table related to Case 2

AGE years	PROB	LEVEL
62	0,15	very-low
64	0,286	low
66	0,4074	low-middle
68	0,5141	middle
70	0,6064	middle-high
72	0,6851	middle-high
74	0,7512	high
76	0,8059	high
78	0,8505	high
80	0,8864	high

The prediction engine has also produced the qualitative view represented by the histogram showed in Fig. 4.

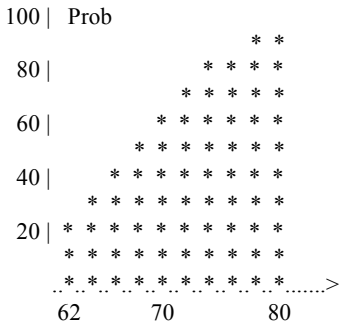

Fig. 4 Predictions histogram related to Case 2

The possibility of simulating, and then compare, the future effects of possible alternative situations is a powerful facility that CHEERUP provides its users with. Let us consider, for example the following case. Let us suppose that at the first session (age = 60 years) a subjects presents the following situation: Obesity= no, Hypertension= yes, Cholesterol= ok, Smoke= yes *and such state has lasted for* 20 *years*.

3. Case 3

Let us simulate the case in which the subject keeps on being hypertensive and smoker even in the future. By using the *initial* statistical values the prediction engine produces the numeric outcome showed in Table III.

The prediction engine has also produced the qualitative view represented by the histogram showed in Fig. 5.

Table III Predictions table related to Case 3

AGE years	PROB	LEVEL
62	0,09	very-low
64	0,181	low
66	0,2793	low
68	0,373	low-middle
70	0,4608	middle
72	0,5471	middle
74	0,6241	middle-high
76	0,6918	middle-high
78	0,7504	high
80	0,8003	high

4. Case 4

Let us simulate the case in which, after the current session, the subject is no more hypertensive and does not smoke any more. The fact that the subject stops smoking after he has smoked for 60 years is represented by selecting the Smoke state: "no but yes in the past (for 15 < TI ≤ 20 years)" *and such situation has lasted for* 0 *years*. By using the *initial* statistical values the prediction engine produces the numeric outcome showed in Table IV.

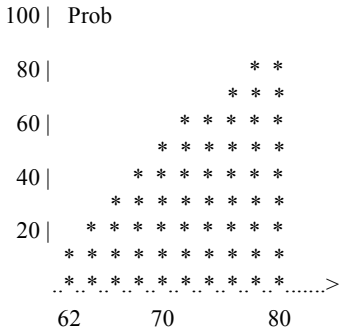

Fig. 5 Predictions histogram related to Case 3

The prediction engine has also produced the qualitative view represented by the histogram showed in Fig. 6.

D. Results analysis and discussion

Table IV Predictions table related to Case 4

AGE years	PROB	LEVEL
62	0,05	very-low
64	0,0785	very-low
66	0,1154	very-low
68	0,1596	low
70	0,21	low
72	0,2574	low
74	0,3094	low-middle
76	0,3646	low-middle
78	0,4218	low-middle
80	0,4796	middle

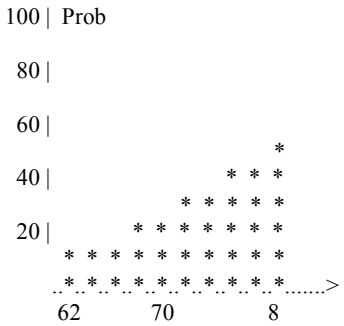

Fig. 6 Predictions histogram related to Case 4

We can easily understand where the numbers come from. Let us consider, for example, the first two rows of Table II.

The first row, corresponding to age 62, shows the number 0.15. It represents L_{62} , the value learned by the prediction engine. In fact (N_{62} | Obesity= ”yes”, Hypertension= ”yes”, Cholesterol= “not-ok”, Smoke= “yes” *and such state has lasted for* 40 *years*) = 15. Such value represents the number of persons (among the initial population of 100 persons) that at the age of 62 years have had the first cardiac infarct. As for the second row of Table II, row corresponding to age 64, we can apply the (9) and state that:

$$P(E_{64} = y \mid E_{60} = n, Obesity_{64} = yes,$$
$$Hypertension_{64} = yes, Cholesterol_{64} = not-ok,$$
$$Smoke_{64} = yes...for...44...years) =$$
$$L_{64} \cdot (1 - X_{62}) + X_{62} = 0.16 \cdot (1 - 0.15) + 0.15 =$$
$$0.286$$

It is interesting to compare Case 4 with Case 1. Let us notice that in Case 4 the number of occurrences of first cardiac infarcts for age = 62 is 5, whereas in Case 1 it is 1. Then, for age = 72, the two cases proceed with the same numbers of occurrence (Table V). In fact, referring to Case 4, it can be stated that at the age of 72 years the subject has not smoked for 12 years. As a consequence, since we have defined the rule (in subsection B) that after a time period greater than 10 years the negative effects of smoking (regarding cardiac infarct) become null, from the age 72 on the occurrence numbers in Case 4 increase with the same rate than in Case 1. Notwithstanding this fact, the probability values remain different in the two cases even after the age of 72 years. This is due to the fact that the probability calculated for an age uses the probability calculated for the preceding age.

Table V Comparison between Case 4 and Case 1

AGE (years)	Occur. number Case 4	Occur. number Case 1
62	5	1
64	3	2
66	4	3
68	5	4
70	6	5
72	6	6
74	7	7
76	8	8
78	9	9
80	10	10

V. RELATED WORK AND DISCUSSIOON

Industry is a typical world in which predictive monitoring, mostly intended as preventive monitoring, has find numerous applications with a variety of approaches. Twenty years ago already, preventive monitoring was a crucial theme for manufacturing processes (typically, for example, in the world of the large car manufacturing companies [1]). In manufacturing industries there is a considerable attention to reduce costly and unexpected breakdowns. As a consequence preventive maintenance is becoming more and more important. Maintenance should abandon the traditional “fail and fix” approach to pass to the more modern “predict and prevent” one [2]. As a consequence the fundamental need is monitoring degradation instead of detecting faults. A predictive performance and degradation monitoring is what is needed for an effective proactive maintenance to prevent machines from breakdown. The theme of degradation monitoring for failure prevention applied to vehicle electronics and sensor systems is faced in [3] where the authors propose a unified monitoring and prognostics approach that prevents failures by analyzing degradation features, driven by physics-of-failure. The need, for manufacturers of complex systems, to optimize equipment performance and reduce costs and unscheduled downtime, gives rise to system health monitoring. System states monitoring is augmented with prediction of future system health states and predictive diagnosis of possible future failure states [4]. Predictive monitoring has been also applied to flexible manufacturing systems. In [5], the main objective is to manage progressive failures in order to avoid breakdown state for the flexible manufacturing system. The approach to predictive monitoring proposed in [6] uses predictions from a dynamic model to predict whether process variables will violate an emergency limit in the future (predictions are based on a Kalman filter and disturbance estimation). Predictive monitoring has also been applied in many specific industry worlds like, for example, press manufacturers [7] and chemical plants [8]. In many industrial applications predictive monitoring assumes the meaning of preventive monitoring and aims to enhance the effectiveness of preventive maintenance by making it proactive. In some cases though, predictive monitoring is finalized to early intervening to maintain a system at a high level of performance. It is the case of a predicting monitoring application for wireless sensor networks: “...by monitoring and subsequently predicting trends on network load or sensor nodes energy levels, the wireless sensor network can proactively initiate self-reconfiguration…” [9]. In most industry applications the acquisition of monitoring data is carried out through sensors [10].

Predictive monitoring has found many applications in medicine too. In general they are specific applications. For example, interesting applications have been carried out in the field of diabetes therapy. In [11] and [12], continuous glucose monitoring devices provide data that are processed by mathematical forecasting models to predict future glucose levels in order to prevent hypo-/hyperglycemic events. Many other specific applications of preventive monitoring may be found in medicine [13], [14].

VI. Conclusion

This section presents the mathematical foundation of the Prediction Engine of CHEERUP. It is organized as follows.

The generality of CHEERUP is due to the great heterogeneity of the portals that it is possible to build and use in various domains.

The simplicity of CHEERUP is due to: friendly user-interface (simplicity in using), modular structure (simplicity inside), correctness and coherence controls (simplicity in assistance), theoretical model underlying prediction (simplicity in theory), explanation of the reason why predictions are what they are (simplicity in comprehension).

The effectiveness of CHEERUP is due to the numerous powerful facilities offered to the user. Let us think, for example, of the possibility of testing a portal in the Portals Using environment before definitely terminating its building process (if testing reveals some imperfections, the portal can still be modified by turning back into the Portals Building environment). Let us think of the possibility of simulating, for the subject under examination, various combinations of conditions (context states) under which the future time elapses, getting, as a consequence, the related future probabilities of E occurrence (this might be useful in case of trade-off problems concerning the best measure to be taken in advance). And so forth.

The efficiency of CHEERUP is due to the fact that every specific combination of context states is dynamically created when it is required so to avoid the combinatorial explosion consequent to the creation in advance of all the possible combinations.

Finally, let us conclude by mentioning the CHEERUP structural propensity to favor co-operation among working groups by means of several facilities useful to work in team, in structured organizations.

In order to provide whoever is interested in having a look inside the proposal with the possibility of experiment it and understand it more deeply, CHEERUP is equipped with a self-demo facility (including a user-friendly demo guide), an infrastructure that allows an interested reader to build and test his/her own demo-portal without interfering with real portals possibly present in CHEERUP at the time of the demo.

The interested reader can find CHEERUP at the Web-address: www.cheerup.it

References

[1] S. Spiewak and M. Szafarczyk, "A Predictive Monitoring and Diagnosis System for Manufacturing", *CIRP Annals - Manufacturing Technology,* vol. 40, no. 1, pp. 400-403, 1991.

[2] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu and H. Liao, "Intelligent prognostics tools and e-maintenance", *Computers in Industry*, vol.57, no. 6, pp. 476-489, 2006. [Online]. Available: http://dx.doi.org/10.1016/j.compind.2006.02.014

[3] H. Liao and J. Lee, "Predictive Monitoring and Failure Prevention of Vehicle Electronic Components and Sensor Systems", *SAE 2006 World Congress & Exhibition, April 2006*, Detroit, MI, USA, Session: Automobile Electronics and Systems Reliability (Part 1 of 2).

[4] R. Kothamasu, S. H. Huang and William H. VerDuin, "System health monitoring and prognostics—a review of current paradigms and practices", *The International Journal of Advanced Manufacturing Technology*, vol. 28, no. 9-10, pp. 1012-1024, 2006. [Online]. Available: http://dx.doi.org/10.1007/s00170-004-2131-6

[5] F. Ly, A. K. A. Toguyeni and E. Craye, "Indirect predictive monitoring in flexible manufacturing systems", *Robotics and Computer-Integrated Manufacturing*, vol. 16, no. 5, pp. 321-338, 2000. [Online]. Available: http://dx.doi.org/10.1016/S0736-5845(00)00015-6

[6] B. C. Juricek, D. E. Seborg and W. E. Larimore, "Predictive monitoring for abnormal situation management", *Journal of Process Control*, vol. 11, no. 2, pp. 111-128, 2001. [Online]. Available: http://dx.doi.org/10.1016/S0959-1524(00)00043-3

[7] S. A. Spiewak, R. Duggirala and K. Barnett, "Predictive Monitoring and Control of the Cold Extrusion Process", *CIRP Annals - Manufacturing Technology*, vol. 49, no. 1, pp. 383-386, 2000.

[8] J Jeng, C Li, H Huang, "Dynamic Processes Monitoring Using Predictive PCA", *Journal of the Chinese Institute of Engineers*, vol. 29, no. 2, pp. 311-318, 2006.

[9] A. Ali, A. Khelil, F. K. Shaikh and N. Suri, "MPM: Map based Predictive Monitoring for Wireless Sensor Networks", presented at Autonomic Computing and Communications Systems, Third Int. ICST Conf. Autonomics 2009, Limassol, Cyprus, September 9-11, 2009.

[10] S. C. Choi and P. A. Pepple, "Monitoring Clinical Trials Based on Predictive Probability of Significance", *Biometrics*, vol. 45, no. 1, pp. 317-323, 1989. [Online]. Available: http://www.jstor.org/stable/2532056

[11] J. Reifman, S. Rajaraman, A. Gribok and W. K. Ward, "Predictive Monitoring for Improved Management of Glucose Levels", *Diabetes Science Technology*, vol. 1, no. 4, pp. 478-486, 2007.

[12] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, C. Cobelli, E.J. Gómez, M. Rigla, A. de Leiva and M.E. Hernando, "Artificial Neural Network Algorithm for Online Glucose Prediction from Continuous Glucose Monitoring", *Diabetes Technology & Therapeutics*, vol. 12, no. 1, pp. 81-88, 2010. [Online]. Available: http://dx.doi.org/10.1089/dia.2009.0076

[13] J. Chen, T.-Y. Hsu, C.-C. Chen, and Y.-C. Cheng, "Online Predictive Monitoring Using Dynamic Imaging of Furnaces with the Combinational Method of Multiway Principal Component Analysis and Hidden Markov Model", *Industrial & Engineering Chemistry Research*, vol. 50, no 5, pp. 2946-2958, 2011. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/ie100671j

[14] D. P. O'leary, L. L. Davis and S. Li, "Predictive Monitoring of High-frequency Vestibulo-ocular Reflex Rehabilitation Following Gentamicin Ototoxicity", *Acta Oto-Laryngologica*, vol. 115, no. S520, pp. 202-204, 1995.

[15] F. V. Jensen, *An Introduction to Bayesian networks*, London: UCL Press, 1996.

Silvano Mussi received the degree in Physics in 1975 from the University of Milan, Italy. He has been with a telecommunication company (ITALTEL) for over 10 years, working in the Research and Development department, in the fields of Software Engineering and Functional Discrete Simulation of Real-time Systems.

Then, for over 20 years he has been with an interuniversity consortium for information technology (CILEA) where for over 10 years he has played the role of marketing manager and for over 10 years he has done research in the fields of Knowledge Engineering and Expert Systems, co-operating with University of Brescia and Milan Polytechnic. For three academic years he has been contract-professor of Artificial Intelligence at the University of Brescia.

His current research interests are in the fields of probabilistic expert systems, predictive monitoring and decision support systems.

His scientific publications can be found at the Web-page: www.cheerup.it/CHEERUP/publications.html.