Automated Data Validation and Repair Based on Temporal Ontologies

Silvia Miksch[1)], Werner Horn[1,2)], Christian Popow[3)], Franz Paky[4)]

[1)]Austrian Research Institute for Artificial Intelligence (ÖFAI)
Schottengasse 3, A-1010 Vienna, Austria
phone: ++43 - 1 - 53 53 281 - 0
fax: ++ 43 -1 - 53 20 652
email: silvia@ai.univie.ac.at

[2)] Department of Medical Cybernetics and Artificial Intelligence, University of Vienna
Freyung 6, A-1010 Vienna, Austria
phone: ++43 - 1 - 53 53 281 - 0
fax: ++ 43 -1 - 53 20 652
email: werner@ai.univie.ac.at

[3)] NICU, Division of Neonatology, Department of Pediatrics, University of Vienna
Währinger Gürtel 18-20, A-1090 Vienna, Austria
phone: ++43 - 1 - 40 400 - 32 32
fax: ++ 43 -1 - 40 400 - 31 65
email: christian.popow@vm.akh-wien.ac.at

[4)] Department of Pediatrics, Hospital of Mödling
Weyprechtgasse 12, A-2340 Mödling, Austria
phone: ++43 - 2236 - 204 - 0
fax: ++ 43 - 1 - 2236 - 204 - 111

Abstract

Most knowledge-based monitoring and therapy planning systems neglect the importance of data validation. Real data are more faulty than expected. Moreover, only reliable data may be used for effective and efficient therapy planning. Additionally, most systems do not take into account the various kinds of data and the various frequencies at which they are usually available.

We propose automated data validation methods which consider the various kinds of real data and which are based on temporal ontologies (time points, time intervals, and trends) in order to arrive at reliable data. Furthermore, our approach includes repair and adjustment methods for correcting wrong or ambiguous data. Our approach benefits from dynamically derived qualitative data-point- and trend-categories which result in unified qualitative descriptions of parameters and overcome the limitations of comparison with predefined static thresholds.

Our methods are applicable to domains where different kinds of data are available and where no reliable structure-function model exists because the underlying mechanism is only poorly understood. We applied them in VIE-VENT, an open-loop knowledge-based system for artificially ventilated newborn infants.

Keywords:

Reasoning about Action and Change
Principles of AI Applications
Knowledge-based Monitoring and Therapy Planning (Temporal Reasoning),
Artificial Ventilation

Declaration:

This Paper has not already been accepted by and is not currently under review for a journal or another conference. Nor will it be submitted for such during IJCAI's review period.

1. Introduction

This work is part of the growing field of monitoring and therapy planning in medical domains. We were particularly motivated by the real-world problems of such processes facing an unexpectable high amount of faulty data and various types of data available occurring at various frequencies (e.g., high or low frequency data). Moreover, there exists no reliable structure-function model because the underlying mechanism is often poorly understood.

Our approach is oriented on, but not limited to, our application domain: artificial ventilation of newborn infants. The technical improvement of Intensive Care Units' (ICUs') equipment makes a huge amount of data available to the medical staff, but methods for data validation to arrive at reliable measurements are missing. Several monitors have a built-in module for recognizing unusual data values, especially those arising from hardware problems. But these built-in modules often trigger false alarms. The monitoring data (signals) are observed by the trained medical staff. However, these single observations are only recognized for being "normal" or "abnormal". Information about trends, "natural" oscillations, etc. is very difficult to gather. Inexperienced personnel may have difficulties in interpreting a clinical picture from single monitoring data. Some of the variables are influenced by other clinical variables that may not be (continuously) determined (like cardiac output, pulmonary perfusion). Physicians recognize many respiratory factors, like perfusion or oxygenation, but cannot predict quantitative effects on the blood gas measurements from changes of these factors. Therefore in this domain time-series analysis techniques (Avent and Charton, 1990), may be insufficient because of the absence of an appropriate curve-fitting model.

Our approach was to develop different methods for data validation and repair algorithms which are based on different temporal ontologies (Allen, 1991; Dean and McDermott, 1987) due to a real-world monitoring situation. The utility of our approach is illustrated by VIE-VENT, an open-loop knowledge-based system for artificially ventilated newborn infants. Our aim in developing VIE-VENT (Miksch, et al. 1993) was to incorporate alarming, monitoring, and therapy planning tasks within one system in order to overcome some of the limitations of existing systems (e.g., GUARDIAN (Hayes-Roth, et al. 1992; NeoGanesh/Ganesh (Dojat, et al. 1992, 1994)). VIE-VENT is especially designed for practical use under real-time constraints at neonatal ICUs (NICUs) and the various components are built in analogy to the clinical reasoning process. The data-driven architecture of VIE-VENT consists of several modules: data selection, data validation, data abstraction, data interpretation and therapy planning. All these steps are involved in a single cycle of data collection from monitors. VIE-VENT's knowledge-base is implemented in Clips (v6.02, COSMIC/NASA), a forward chaining rule and/or object based development system. All examples used are defined in Clips notation.

In the first part we categorize the possible kinds of data that are available during the patient monitoring processes. The second part focuses on different methods for data validation based on different temporal ontologies and their reasoning process. Additionally, we present repair and adjustment methods for correcting wrong or ambiguous data.

2. Kinds of data

The data used are divided in the continuously and discontinuously assessed monitoring data and their corresponding derived qualitative data-point- and trend-categories.

2.1 Original monitoring data

According to the technical equipment of modern ICUs a huge amount of on-line data is available. Additionally off-line data and qualitative observations are available and needed for a global picture of the patient's condition and for an effective reasoning process. The data can be divided according to their observation frequency and regularity as well as for their data types. We distinguish between three kinds of data: continuously assessed quantitative data, discontinuously assessed quantitative, and qualitative data.

The most important kind of data in the field of monitoring and therapy planning are continuously observed quantitative measurements which are successively and regularly taken real numbers (e.g., transcutaneous blood gas measurements $P_{tc}CO_2$, $P_{tc}O_2$, S_aO_2). Usually series of different parameters are monitored and compared at once. The discontinuously assessed quantitative data are not regularly taken real numbers which are used in critical situations as control factors for cross validation (e.g., dynamic calibration) and therapy planning (e.g., invasive blood gas measurements P_aCO_2, P_aO_2, pH). Discontinuously assessed qualitative data are not regularly taken verbal descriptions which are also used for cross validation and therapy planning (e.g., excessive chest wall expansion, low spontaneous breathing effort). The discontinuous data are entered on request at a particular time-point, but may be valid for a longer time period.

2.2 Dynamically derived qualitative data-point- and trend-categories

In addition to the numerical time-point- and trend-values we derive qualitative data-point- and trend-categories to detect faulty measurements. The aim of this data abstraction process is to arrive at unified qualitative descriptions of data points and trend data. It transforms quantitative measurements into qualitative values, which can be used in the system model for data interpretation and therapy planning. An advantage of using qualitative values is their unified usability in the system model, no matter of which origin they are. Adaptation to specific situations can easily be done by using specific transformation tables without changing the model of data interpretation and therapy planning. Additionally, by using qualitative values an easily comprehensible and transparent system model can be developed.

(a) Qualitative *data-point*-categories

The transformation of quantitative data points into qualitative values is usually performed by dividing the numerical range of a parameter into regions of interest. Each region stands for a qualitative value. The region defines the only common property of the numerical and qualitative values. It is comparable to "point temporal abstraction" of (Shahar and Musen, 1992; Shahar, 1994) .

The basis of the transformation of the blood gas measurements are *data-point-transformation schemata* relating single values to seven qualitative categories of blood gas abnormalities (qualitative *data-point*-categories).These data-point-transformation schemata are defined for all kinds of blood gas measurements depending

on the blood gas sampling site (arterial, capillary, venous, transcutaneous) and the mode of ventilation (IPPV, IMV). The different modes of ventilation require specific predefined target values depending on different attainable goals. Figure 1 shows the scheme of $P_{tc}O_2$ during IPPV. For example, the transformation of the transcutaneous $P_{tc}O_2$ value of 91 mmHg during IPPV results in a qualitative $P_{tc}O_2$ value of *g3* ("extremely above target range"). The $w_{i,x}$ -values divide the qualitative regions. The transformation of trends is based on these qualitative data-point-categories, which are described in the following section.

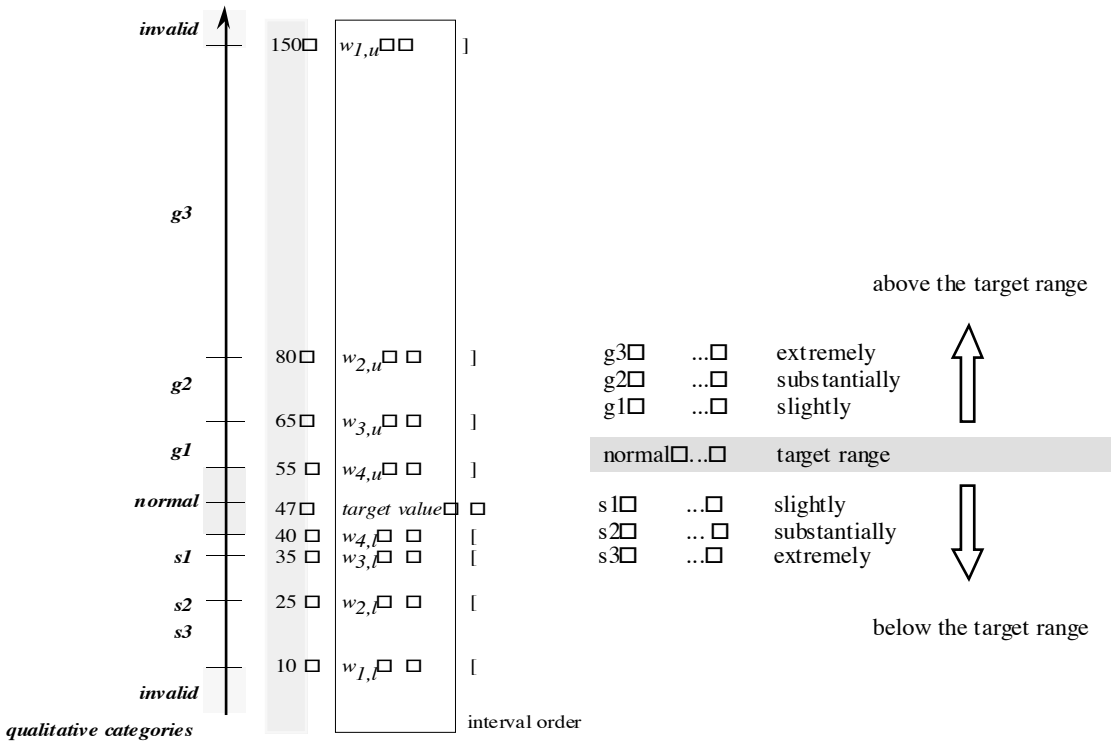

Figure 1: Data-point-transformation scheme of $P_{tc}O_2$ during IPPV

(b) Qualitative *trend*-categories

The transformation of trend data into qualitative values is based on the combination of qualitative data-point-categories and the qualitative descriptions of the expected behavior of a parameter (*expected qualitative trend description*s; e.g., "parameter $P_{tc}O_2$ is moving one qualitative step towards the target range within 10 to 30 minutes"). These *trend-curve-fitting schemata* transform the quantitative trend values into ten qualitative categories guided by physiological criteria (figure 2). We used a dynamic comparison algorithm to classify the trend data, which performs a stepwise linearization of the expected exponential function to overcome complexity (compare Miksch, et al. 1994b).

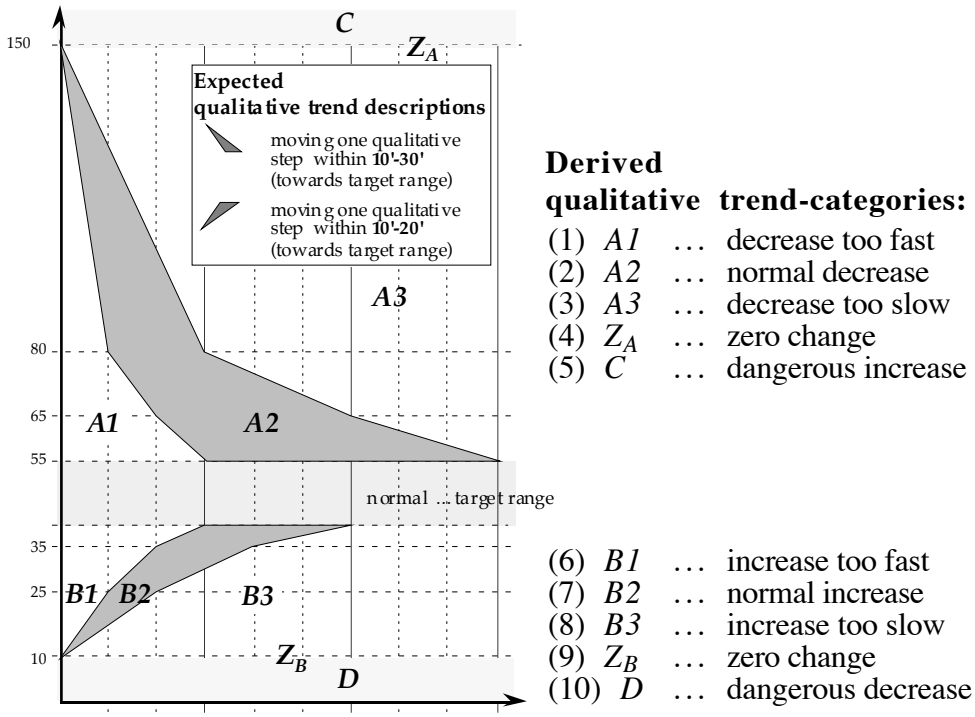

Figure 2: Trend-curve-fitting scheme

The aim of our approach is to use all available kinds of data for data validation based on different temporal ontologies.

3. Definitions and overview

The major aim of the data validation process is to detect faulty measurements or artifacts and finally to arrive at reliable measurements which may be used for further analysis tasks. An artifact is a situation where a measured variable does not reflect the clinical context. Undoubtedly, the data validation is an important, but often neglected part of the monitoring and therapy planning process.

For definitional purposes, it is essential to distinguish between our usage of the term "(data) validation" and "validation of knowledge-based systems". The latter is the process which attempts to determine whether a system does or does not satisfy one of its specifications. In this sense validation is the sum of verification (the proof of objective, formal specifications) and evaluation (the proof of interpretative, pseudo-formal specifications) (Laurent, 1992).

We deal with "data validation". It is the context-sensitive examination of the plausibility of input data based on different temporal ontologies. The result is a classification of the input data. An input value is finally classified as

(a) correct
(b) wrong
(c) unknown
(d) adjusted

A measurement is classified as "adjusted" if a "wrong" or "unknown" value is corrected by a repair or adjustment method (see section 5). If VIE-VENT recognizes a faulty measurement but could not apply any of these methods, it is classified as "wrong". If VIE-VENT receives no data for a measurement from the monitor and no value could be estimated, then the measurement is classified as "unknown". Otherwise it is classified as "correct". Not all methods mentioned below lead to a final classification. Some of them (like, the time-point-based functional depen-

dencies) result in an intermediate and ambiguous classification of "some are wrong". This information is forwarded to and handled by the repair and adjustment module which provides strategies for repairing and adjusting not plausible or missing values based on the same temporal ontologies as the data validation module.

We divided our methods in three parts based on their underlying temporal ontologies: time-point-based, time-interval-based, and trend-based reasoning. Figure 3 gives an overview of the particular categories.

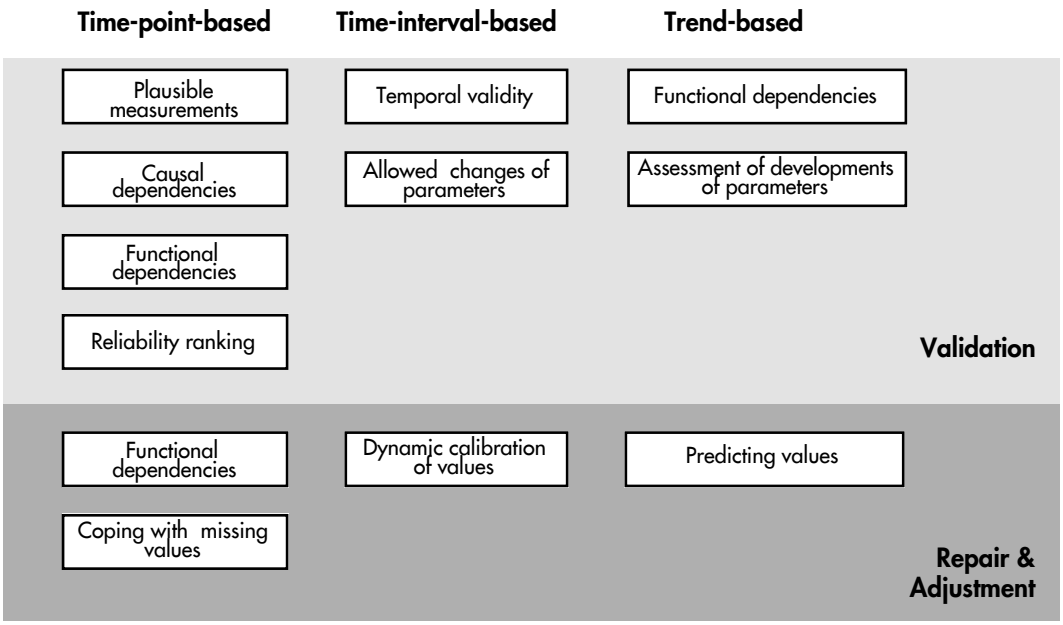

Figure 3: Overview of the components of the data validation and repair/adjustment modules

4. Data validation

4.1 Reasoning based on time points

The time-point-based concept uses the value of a parameter at a particular time point for its reasoning process. It benefits from the transparent and fast to proceed reasoning process. But it suffers from neglecting any information about the history of the observed parameters. VIE-VENT uses the following methods for detecting faulty measurements or artifacts: checking the plausibility of measurements, causal and functional dependencies, and reliability ranking. Parts of the time-point-based concept were discussed in a previous publication (Miksch, et al. 1994a).

4.1.1 Plausible measurements

The most basic method is time-point based range checking. We have enhanced this method by adding additional attributes, which define the clinical context (e.g. arterial, IPPV). There are look-up tables for all input parameters covering the plausible ranges. A parameter in the look-up table is specified by a parameter name, a list of attribute descriptors, an upper limit and a lower limit. For example, (pCO_2, (arterial, IPPV), 10, 140), where “arterial" refers to the kind of blood gas analysis and IPPV to the mode of ventilation. When a new parameter value is received, the system checks if this value is within or out of range and a corresponding flag ("correct" or "wrong") is set, e.g., if $10 \leq pCO_2(\text{arterial, IPPV}) \leq 140$ then it is a "correct" measurement.

4.1.2 Causal dependencies

A causal dependency specifies a relationship between an actual value and the expected value of a corresponding parameter. It is a kind of cross-relation between different measurements. For example, if the chest wall expansion is small then the tidal volume has to be less-than-or-equal-to (≤) 5 ml/kg. This method cannot rate a particular parameter as "correct" or "wrong". If the causal dependency fails all involved parameters are marked as "some are wrong". This information is forwarded to the repair and adjustment methods (compare chapter 5). If this ambiguity could not be solved all values are classified as "wrong".

4.1.3 Functional dependencies

A time-point based functional dependency describes a functional relationship between two or more parameters to check inadequate data transmission or faulty measurements. There are two possible kinds of functional dependencies: the involved parameters refer to the same clinical context or to a different one. In the first case we have to deal with some kind of calibrating measurements. Calibration is a time-point-based method due to its activation at the time of the availability of the data, but its consequences are adjustments of the values for a longer period of time. We categorize these adjustments as a time-interval-based repair method. It is discussed in chapter 5.2.1.

An example of the latter kind is $f = 60 / (t_I + t_E)$, where f is the ventilation rate, t_I is the inspiration time and t_E is the expiration time. We receive all values of f, t_I and t_E and check the functional dependency of these parameters. As in case of causal dependencies, if the functional dependency fails it results in an intermediate, ambiguous classification of "some are wrong" and is handled in the same way.

4.1.4 Reliability ranking

Priority lists of measurements are an indicator of their reliability. The data validation process allows to identify a less reliable parameter from a set of conflicting parameters. The result is a *reliability ranking*. For example, arterial blood gases are more reliable than venous blood gases; P_aO_2 is more reliable than S_aO_2 and S_aO_2 is more reliable than $P_{tc}O_2$. This method is triggered, if an ambiguous classification of values "some are wrong" has been derived.

4.2 Reasoning based on time intervals

The time-interval-based concept deals with values of different parameters during an interval. VIE-VENT uses two different methods: temporal validity and allowed changes of a parameter during an interval.

4.2.1 Temporal validity

The temporal validity sets the valid time intervals of parameters. The discontinuously and continuously assessed data are handled in different ways.

(a) Discontinuous data

The discontinuously assessed quantitative and qualitative data are handled by the same method. There are two possibilities for setting a time interval of a parameter valid.

(1) The user of VIE-VENT can specify a duration of validity when entering a particular discontinuous value.
For example,

"P_aO_2 should be valid for the next 30 minutes" is expressed as

```
(time-interval (kind temporal-validity)
   (parameter PaO2) (validation valid)
   (begin-time 12:00:00) (duration 30 minutes)
   (reason entered-by-the-user))
```

(2) For each discontinuous parameter there are predefined default durations of validity . When no duration is entered, VIE-VENT uses the default interval as long as no external invalidation procedure is activated.

The parameter is set invalid, if one of the following conditions becomes true:

(1) the time interval of the parameter has elapsed;

(2) a new value of the discontinuous parameter is available;

(3) an external event enforces to manually set the parameter invalid.

(b) Continuous data

The continuously assessed data are handled in a different way: instead of valid time intervals we define *invalid* time intervals. A continuous value is only set invalid due to external problems (e.g., new application of sensors, calibration of sensors, disconnection of a sensor).

For example,

```
(time-interval (kind temporal-validity) (parameter SaO2)
   (validation invalid) (begin-time 10:00:00)
   (duration 10 minutes)
   (reason changing-position-of-sensors))
```

means, that S_aO_2 will be invalid from 10:00:00 to 10:10:00 due to calibration and changing of the position of the sensor.

4.2.2 Allowed changes of parameters

The checks for allowed changes of parameters is the comparison of the new value with previously assessed values within a predefined time-interval. This method can only be applied to continuously assessed quantitative values. We distinguish between two situations: allowed changes of parameters without a therapeutic action and allowed changes of parameters after a therapeutic action.

(a) Allowed changes of parameters without a therapeutic action

If no therapeutic action has taken place, it is possible to check the changes of a parameter within a particular time interval. The amount of change and the particular interval is defined for each continuous parameter.

For example,

```
(time-interval (kind allowed-changes)
   (parameter SaO2) (validation valid)
   (duration 5 minutes) (amount 10%)
   (reason without-a-therapeutic-action))
```

means, that only a change of less than 10% of the S_aO_2 value is valid within 5 minutes.

(b) Allowed changes of parameters after a therapeutic action

If a therapeutic action has taken place, we would expect a particular parameter to improve towards the normal range after a certain delay-time. The duty of this method is to specify a larger amount of allowed changes of parameters for a particular time interval.

For example,

```
(time-interval (kind allowed-changes) (parameter SaO2)
   (validation valid) (begin-time 11:00:00)
   (duration 10 minutes) (amount 20%)
   (reason after-a-therapeutic-action))
```

means, that a change of less than 20% of the SaO2 value is valid from 11:00:00 to 11:10:00.

4.3 Reasoning based on trends

The trend-based concept tries to analyze the development of a parameter during an interval. A trend is a significant pattern in a sequence of time-ordered data. Therefore the following methods can only handle continuously observed measurements. It benefits from the dynamically derived qualitative trend-categories which overcome the limitations of predefined static thresholds. VIE-VENT uses the following methods: trend-based functional dependencies and assessment procedures of the development of a parameter.

4.3.1 Preconditions

a) Trend approximation

The problem of planning artificial ventilation of newborn infants - as in other medical fields, like pediatric growth (Haimowitz and Kohane 1993) - lies in the lack of an appropriate curve-fitting model to predict the development of a physiological variable from actual measurements. Therefore our first effort is to approximate the growth of the continuously assessed measurements $P_{tc}O_2$, $P_{tc}CO_2$ and S_aO_2 using a simple linear regression model $(E(Y) = a + k * X_i$, where $E(Y)$ is the expected value, X_i are the observed data points, a is a constant value (offset), and k is growth rate).

Choosing this simple linear regression model was influenced by practical clinical reasons: the only important characteristics of parameters used by physicians are on the one hand increases, decreases, or zero changes of parameters, and on the other hand too slow, too fast, or reasonable changes of parameters. Therefore it would be superfluous to calculate a curve-fitting model of higher order with additional features for our purpose.

We distinguished four kinds of trends based on our samples, which are derived from new measurements every 10 seconds. The distinction of the trends are guided by physiological criteria:

(1) *very short-term* trend: sample of data points based on the *last* minute
(2) *short-term* trend: sample of data points based on the *last 10* minutes
(3) *medium-term* trend: sample of data points based on the *last 30* minutes
(4) *long-term* trend: sample of data points based on the *last 3* hours

The very short-term, the short-term, and the medium-term trend are used for data validation. The long-term trend is used for changing therapeutic strategies.

b) Criteria of validity to calculate trends

During a monitoring process the position of a measurement in the sequence of time-ordered data influences the reasoning process: namely, the recent measurements are more important than the historical measurements. Therefore criteria dealing only with the average distribution of the measurements are insufficient. Due to this precondition we defined two criteria of validity to make sure that the used trend is meaningful: an at least amount of valid measurements within the whole time interval, and an at least amount of valid measurements at the end of the time interval. These limits are defined by experts based on their clinical experience. They may easily be adapted to the situation of a specific clinic based on the frequency at which data values arrive. Table 1 shows the different conditions.

Kind of trend	**valid measurements** *within* **the whole time interval**	**valid measurements** *at the end* **of the time interval**
very short-term trend	at least 30 percent	last 20 percent
short-term trend	at least 40 percent	last 30 percent
medium-term trend	at least 45 percent	last 35 percent
long-term trend	at least 50 percent	last 40 percent

Table 1: Criteria of validity

4.3.2 Trend-based functional dependencies

We can also define functional dependencies for expectations on trends. The increase/decrease of a parameter suggests an increase/decrease of another. If such an expectation is violated, one of these parameters must be faulty (classified as "some are wrong"). The short-term trend is used for comparison.

E.g., if the minute ventilation (AMV) is increasing then $P_{tc}CO_2$ is expected to decrease.

4.3.3 Assessment of developments of parameters

The assessment procedure of developments of a particular parameter is based on the qualitative trend-categories (chapter 2.2) and the ordering of these categories. The preconditions to proceed are a positive judgment of the criteria of validity to calculate trends (chapter 4.3.1) and a valid allowed change of the parameter (chapter 4.2.2). The simple method of comparing the actual numerical growth rate with a predefined growth threshold is not applicable in our case. The reason lies in the physiological behavior which results in different expected normal growth rates depending on the absolute values of the parameter. Therefore we use the ten

qualitative trend-categories which reflect this dynamic aspect. The qualitative trend-categories are divided by the normal region in an upper and a lower region. According to these regions the ordering of the qualitative categories is defined as follows

ordering of the qualitative trend-categories:

qualitative *upper* region: A1 - A2 - A3 - ZA - C

qualitative *lower* region: B1 - B2 - B3 - ZB - D

The assessment procedure compares the previous qualitative short-term trend-category with the actual qualitative short-term trend-category. If the actual category belongs to the same qualitative category or to a qualitative neighboring category of the previous category then the parameter is validated as "correct". Otherwise the parameter is classified as "wrong". We can only apply the short-term trends, because the very short-term trend reacts too rapidly to small oscillations of the values (compare figure 4) and the medium- and long-term trend are too insensitive.

Figure 4 illustrates an example of the assessment procedure of the S_aO_2 (arterial oxygen saturation) time series. At 17:17:00 the S_aO_2 value is classified as "wrong", because the actual qualitative short-term trend-category "C" does not belong to "D" (the same qualitative category as the previous values) or to "ZB" (the neighboring category of the previous categories). The very short-term trend categories are not usable for this purpose, because they show too rapid changes of the qualitative categories (at 17:06:00, 17:09:00, 17:12:00, and 17:14:00).

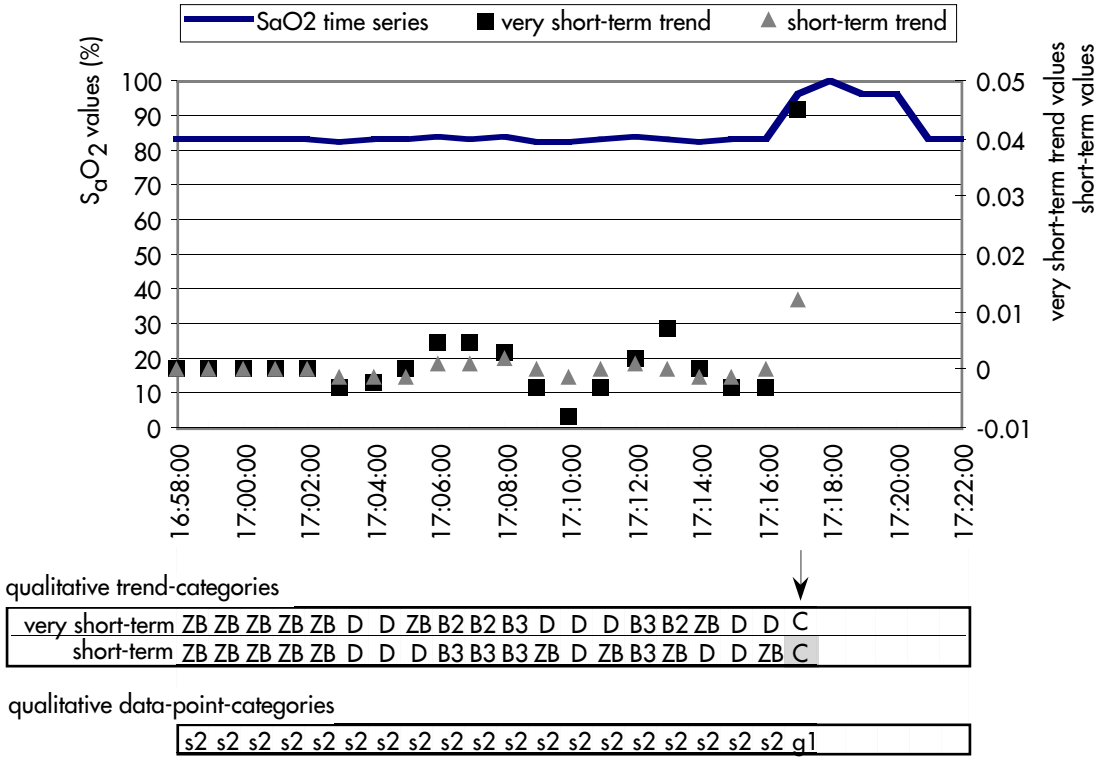

Figure 4: Example assessing the developments of the parameter S_aO_2

4.2 Interaction of the methods

The reasoning methods based on time points and time intervals represent a preprocessing for the reasoning based on trends. They primarily perform static data validation which delivers the necessary preconditions to proceed with the trend-based validation. Further, trend-based validation may result in the conclusion that a

data value of the last point is implausible and has to be invalidated (in some severe cases even older values have to be invalidated, which results in the problem of "the revision of the past"). In such a case previous validation methods have to be reapplied. Given such strong interaction of the methods presented the data validation process has been implemented as a multi-step procedure.

5. Repair and adjustment

If a measurement was classified as "wrong", "some are wrong" or "unknown", VIE-VENT tries to substitute a correct value. Additionally, if several monitor parameters, which should reflect the same clinical context, deviate from each other a dynamic calibration of these values is necessary. There are four repair possibilities: functional dependencies, coping with missing values, dynamic calibration, predicting a "correct" value. The user of VIE-VENT will be informed about the repair action of a particular value and has the opportunity to disable this feature. All the quantitative continuously and discontinuously assessed parameters can be repaired.

5.1 Time-point-based repair and default values

5.1.1 Applying functional dependencies

Functional dependencies are based on time-point values. They are used for the repair of the ventilator settings. Only values which had been classified as "some are wrong" by the time-point-based causal and functional dependencies are involved in this task. The set of dependencies are the same as used in chapter 4.1. But this repair method takes the additional information of the reliability ranking into account.

5.1.2 Coping with missing values

This method is triggered if a value is marked as "unknown", "wrong", or "some are wrong" and if it could not be adjusted by any other method. There are two options to deal with missing values:

(a) Simplified reasoning process

The simplified reasoning process uses only a few parameters. VIE-VENT uses a simplified system model of neonatal respiration during the initial phase when the only reliable continuous measurement is S_aO_2. There are restricted reactions to decrease oxygenation depending on the degree of abnormality of the S_aO_2 and the actual tidal volume (V_T). The V_T is estimated here by the extent of the chest wall expansion.

(b) No solution

When all measurements are "unknown", "wrong", or "some are wrong" or a critical situation has arisen in the past, VIE-VENT is unable to find a solution and the recommendations of appropriate treatments are shifted to the physician.

5.2 Time-interval-based adjustment

5.2.1 Dynamic calibration of values

When we observe various monitoring parameters, some of the parameters reflect the same clinical context. If these parameters deviate from each other due to the individual situation of the patient or due to variations in the environmental

conditions under which the sensors operate, we need a dynamic adjustment. This is done under the assumption that the previous data validation task has classified the data as "correct". The method we use is a linear calibration based on the reliability ranking and on an analysis of real clinical data.

The activation of calibration depends on the time-point-based measurement. The calibration is done in case the qualitative data-point-values differ by two qualitative categories.

For example, one of the benefits of VIE-VENT is the opportunity to combine values of the transcutaneous blood gas monitor with the more reliable but only rarely drawn discontinuous arterial blood gas measurements. Therefore transcutaneous measurements have to be calibrated against arterial blood gas measurements, $P_{tc}CO_2 = corr(P_aCO_2)$, where *corr* is a linear correlation function. The analysis of 442 cases with corresponding measurements results in the following correlation: $P_{tc}CO_2 = corr(P_aCO_2) = 2.226 + 1.039\ P_aCO_2$ with $r^2 = 0.705$. If qualitative data-point-value of $corr(P_aCO_2)$ and of the actual measured $P_{tc}CO_{2,\ meas}$ differ by two qualitative categories the activation becomes true and the new calibrated value $P_{tc}CO_{2,\ calc}$ is calculated as follows,

$$P_{tc}CO_2^* = corr(P_aCO_2) = 2.226 + 1.039\ P_aCO_2$$
$$c = P_{tc}CO_2^* - P_{tc}CO_{2,\ meas}$$
$$P_{tc}CO_{2,\ calc} = c + P_{tc}CO_{2,\ meas}.$$

The dynamic calibration lasts as long as the discontinuous measurement is set valid (compare 4.2.1).

5.3 Trend-based repair

5.3.1 Predicting values

During a monitoring process the position of a sensor has to be changed frequently and regularly. Therefore the measurements are often missing. The implicit assumption of missing measurements during such a position change is that they will be steady keeping their previously observed values. However, we may be more clever by propagating our trends.

There are two possibilities to deal with missing measurements. First, a stepwise backward checking of the last reliable value and continuing with this value as long as no other system change is detected. Second, applying the previously explained linear regression model based on the short-term trend to predict a "correct" value. A precondition is the stability of the trend. The stability is assessed applying the qualitative trend-categories. If the medium-term and short-term qualitative trend-categories are identical, the precondition of intrinsic development of the measurements becomes true. The trend-based prediction of a value is a more accurate action, because it takes the history of the values into account. But the criteria of validity to calculate a trend have to be fulfilled to predict a value.

Dealing with trend data and continuously and discontinuously assessed measurements leads to the problem of "the revision of the past". First, we may detect that we have falsely classified previously observed measurements as correct or wrong respectively (e.g., wrong prediction of a measurement based on calculation of the trend). Second, there exists a delay-time when collecting discontinuous laboratory data. When the data become available they reflect the clinical situation some minutes ago (e.g., at the time the blood sample was drawn). We time-stamp such data and

use this information especially for the calibration process. However, this may make necessary to revise decisions already taken. In real time we cannot withdraw therapeutic actions and their consequences. Instead we have to adjust therapeutic recommendations. We are aware of circumstances which enforce a revision of such past decisions. Future research will take into account such revisions especially during the evaluation whether a therapeutic action was successful or unsuccessful for the patient.

6. Conclusion and further enhancement

We demonstrate methods for automated data validation and repair based on temporal ontologies (time points, time intervals, and trends). They take into account the various types of data available occurring at various frequencies and combine and integrate a bunch of methods for data validation in a real-time environment. It is important to use all available information for data validation, to cross-validate continuously and discontinuously observed data, and to cross-validate data from different sources. Of essential importance is the reliability ranking of data values to reach meaningful conclusions in conflicting situations. Such reliability may result from a priori definitions, from experience, or from dynamic evaluation of the current data set.

Our approach benefits from dynamically derived qualitative data-point- and trend-categories which result in unified qualitative descriptions of parameters and overcome the limitations of comparison with predefined static thresholds.

An important research topic for the future is the integration of methods which learn from past experience, both from data collected from an individual patient in the past, and from similar situations in our patient data base. Additionally, we will try to integrate strategies solving the problem of "the revision of the past".

Acknowledgment

The current phase of the project is supported by the "Jubiläumsfonds der Oesterreichischen Nationalbank", Vienna, Austria, project number 4666. We also greatly appreciate the support given to the Austrian Research Institute of Artificial Intelligence (ÖFAI) by the Austrian Federal Ministry of Science and Research, Vienna.

References

Allen J.F.: Time and Time Again: The Many Ways to Represent Time, *International Journal of Intelligent Systems*, Vol.6, 341-55, 1991.

Avent R.K., Charlton J.D.: A Critical Review of Trend-Detection Methologies for Biomedical Monitoring Systems, *Critical Reviews in Biomedical Engineering*, 17(6), pp.621-59, 1990.

Dean T.L., McDermott D.V.: Temporal Data Base Management, *Artificial Intelligence*, 32(1), 1-55, 1987

Dojat M., Brochard L., Lemaire E., Harf A.: A Knowledge-Based System for Assisted Ventilation of Patients in Intensive Care Units, *International Journal of Clinical Monitoring and Computing,* 9, pp.239-50, 1992.

Dojat M., Sayettat C.: Aggregation and Forgetting: Two Key Mechanisms for Across-Time Reasoning in Patient Monitoring, in Kohane I.S., et al.(eds.), *AI*

in Medicine: Interpreting Clinical Data, AAAI Press, Menlo Park, pp.33-36, 1994.

Haimowitz I.J., Kohane I.S.: Automated Trend Detection with Alternate Temporal Hypotheses, in Bajcsy R.(ed.), *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, Morgan Kaufmann, San Mateo, CA, pp.146-151, 1993.

Hayes-Roth B., Washington R., Ash D., Hewett M., Collinot A., Vina A., Seiver A.: Guardian: A Prototype Intelligent Agent for Intensive-Care Monitoring, *Artificial Intelligence in Medicine*, 4(2), pp. 165-66, 1992.

Laurent J.-P.: Proposals for a Valid Terminology in KBS Validation, in Neumann□B.(ed.), *Proceedings of the Tenth European Conference on Artificial Intelligence (ECAI92)*, Wiley, Chichester, UK, pp.829-834, 1992.

Miksch S., Horn W., Popow C., Paky F.: Context-Sensitive Data Validation and Data Abstraction for Knowledge-Based Monitoring, *Proceedings of the 11th European Conference on Artificial Intelligence (ECAI 94)*, John Wiley, 1994a.

Miksch S., Horn W., Popow C., Paky F.: Therapy Planning Using Qualitative Trend Descriptions, Österreichisches Forschungsinstitut für Artificial Intelligence, Wien, TR-94-31, 1994b.

Miksch S., Horn W., Popow C., Paky F.: VIE-VENT: Knowledge-Based Monitoring and Therapy Planning of the Artificial Ventilation of Newborn Infants, Andreassen S., et al. (eds.): *Artificial Intelligence in Medicine: Proceedings of the 4th Conference on Artificial Intelligence in Medicine Europe (AIME-93)*, IOS Press, Amsterdam, pp.218-29, 1993.

Shahar Y., Tu S.W., Musen M.A.: Knowledge Acquisition for Temporal Abstraction Mechanisms, Special Issue: Knowledge Acquisition for Therapy-Planning Tasks, *Knowledge Acquisition*, 4(2), 1992.

Shahar Y.: A Knowledge-Based Method for Temporal Abstraction of Clinical Data, *Thesis*, Department of Computer Science, Stanford University, 1994.