Managing Multiple Hypotheses with Agents to Handle Incomplete and Uncertain Data
Benoît Vettier, Laure Amate, Catherine Garbay, Julie Fontecave-Jallon, Pierre Baconnier

To cite this version:
Benoît Vettier, Laure Amate, Catherine Garbay, Julie Fontecave-Jallon, Pierre Baconnier. Managing Multiple Hypotheses with Agents to Handle Incomplete and Uncertain Data. URMASSN’11 - First workshop on Uncertainty Reasoning and Multi-Agent Systems for Sensor Networks, Jun 2011, Belfast, Northern Ireland, United Kingdom. 12p., 2011. <hal-00742112>
Managing Multiple Hypotheses with Agents to Handle Incomplete and Uncertain Data

Benoît Vettier\textsuperscript{1,2}, Laure Amate\textsuperscript{1,2}, Catherine Garbay\textsuperscript{1,2}, Julie Fontecave-Jallon\textsuperscript{1,3}, and Pierre Baconnier\textsuperscript{1,3}

\textsuperscript{1} UJF Grenoble-1 / CNRS, Université de Grenoble, France
\textsuperscript{2} LIG - UMR5217, AMA team
\textsuperscript{3} TIMC - UMR5525, PRETA team

Abstract. In this paper, we study health monitoring, using ambulatory sensors, where the data available are limited, and can be both unreliable and ambiguous. Hence, the need to consider a person’s context: surrounding environment and previous situations. We propose studying multiple situational hypotheses, and the relations between hypotheses present and past. Such hypotheses are managed with a multi-agent system: the agents embody hypotheses on several levels of abstraction, from a general, rough scenario, down to precise states of both physiology and activity. These agents’ hypotheses are evaluated and compared so that plausible hypotheses emerge. We discuss both the representation of situations, and multi-agent adaptive control mechanisms. This is a mainly theoretical approach, although these proposals are illustrated by an application on real data from a daily office life scenario.

1 Introduction

Our goal is to design a system to monitor a person’s situation, using data from ambulatory sensors and expectations regarding a planned activity, in order to detect any alarming event or pattern thereof. The system is to be tailored to this person’s specific physiology, and adaptable to the environment and the general context. For this, we need to bridge a semantic gap between noisy, incomplete and unreliable data from the sensors, and a set of loosely-defined situation models. However, a set of physiological data is only “alarming” depending on the context (for example: a high heart rate is normal if the subject is running), and the sensor data are both unreliable and ambiguous: several different situations can correspond to a given set of observed data.

In this paper, we propose a layered representation of context with the concepts of \textit{State} and \textit{Micro-scenario}. We handle uncertainty with \textit{multiple hypotheses} on these levels, each managed by an autonomous agent, with heterogeneous and personalized data models, and using a mechanism of \textit{prediction-verification}. These hypotheses/agents are created, evaluated, and destroyed according to their plausibility with regards to the observed data and the global population, for which we also introduce \textit{probe} agents to monitor and regulate the agent population, based on a dynamically constructed contextual frame. This work’s topic of
research, which was partly funded by the *Superco* project of the DGA (French defense procurement agency), is therefore both the representation of situations, and the management of multiple and varying hypotheses by a multi-agent system; it does not consider hardware issues (such as sensor communication).

### 2 State of the Art

In this paper, we consider the complex problem of automated medical situation recognition, from an array of physiological and activity sensors, which are worn by a person. Bridging the semantic gap between the sensor data and the *situation* is often seen as requiring a division into smaller sub-problems which are not, by themselves, enough to reason globally [AD03].

Adding to the complexity is the fact that the data from these sensors can be unreliable [SD05], and in the case of human physiology, ambiguous: a set of observations can correspond to several possible situations [RP07]. As a sensor’s reliability can be known a-priori depending on the context, one needs both a structured method of representing context, and a means of adaptive control on the overall sensor-reasoning link.

Context can be seen as scripted scenarios [Cr06], composed of several layers of abstractions, formally linked by relations and roles in situation networks. This is handled with autonomous, auto-regulated modules using prior knowledge, separated from the algorithms [AB05], and handling information about the system’s past and present state.

Adaptive control can be considered as selecting the most relevant algorithms (a selection which must be learnt [PQ07]), tweaking local parameters and changing decision plans [MG03], dynamically focusing the system’s resources on the most informative and/or crucial sensors [MM08][Ha95], and choosing between anticipating long-term problems and fixing immediate issues [AP10].

To handle the ambiguity, and the fact that situation models are defined heterogeneously with regards to the sensors, we chose to manage multiple hypotheses. These hypotheses can then be evaluated and compared through a confidence score, which has links to both POMDP [TY08] and fuzzy logic [PB10]. [PB10] seeks to provide “embedded decision support”, in the form of adaptive alarms, based on high level information extraction and intelligent information processing.

### 3 Representing a Person’s Situation

In this section, we describe our layered situation model, and the links between situational hypotheses in the knowledge base. Then, we give an insight on the *alarming situations* we aim at detecting.
3.1 Interpretation and Abstraction Levels

At the lowest abstraction level are the sensor data (Breath Rate $BR$, Heart Rate $HR$, 3-axis accelerometer...). This is tangible evidence of the subject’s health and activity, the ground context.

On the other hand, there is a theoretical context: a planned activity, as a succession of steps: the Scenario; in the project’s military application, this would be a soldier’s mission plan. It is used as a loose, theoretical frame for our hypothesis generation engine. It represents what should happen in the timeframe we study, and this knowledge is used to suggest new hypotheses. It can also be used to reinforce the likelihood of a hypothesis which would correspond to the expected Scenario. Conversely, we can measure the deviation between the detected situations and planned scenario; deviations could be a cause for alarm.

We introduce two intermediate levels for hypothesis formulation (shown in Fig. 1), in order to bridge the gap between the observed data and this Scenario.

The lower one is the State level, of either physiology or activity (noted $E_\varphi$ and $E_\alpha$, respectively). This level is composed of a limited set of observable models, to interpret the sensor data based on proven medical and actimetric knowledge.

The second level is made of Micro-scenarios ($\mu S$), which are combinations of these States, and represent steps in the global Scenario; the models to interpret the State hypotheses in terms of Micro-Scenarios are based on common sense knowledge or acquired through dedicated learning processes. The $\mu S$ set is, by construction, larger and more open than the State set.

3.2 Prior Knowledge

In the timeframe we study (several hours), a person can be in several successive situations, following (or not) the steps of the Scenario, which is a chain of couples ($\mu S_i$, $t_i$) representing Micro-scenario $\mu S_i$ supposedly beginning at time $t_i$.

Micro-scenarios are a composition of States $E_\varphi$ and $E_\alpha$, associated with contextual information and a meaning in terms of “task knowledge” (as per Fig. 1). Such a composition is of the form:

$$\mu S_x \equiv E_i \land E_j \land \ldots \text{, where } E_i = E_{i1} \lor E_{i2} \lor \ldots$$

Fig. 1: Interpretation/Knowledge Levels. Example of $\mu S_{\{Nap\}}$ composed of $E_\varphi_{\{Drowsiness\}}$ and $E_\alpha_{\{Supine\}}$.  

\[ \text{(1)} \]
meaning:

1. \( \land \): the \( E_i \) hypothesis is a necessary part of this Micro-scenario.
2. \( \lor \): the \( E_i \) hypothesis is among a group of possibly coexisting or successive possible States in the Micro-scenario.
3. \( E_i \) is either a state \( E_\varphi/E_\alpha \), or a group of possible such states.

For example, a Desk work Micro-scenario could be:

\[
\mu S\{\text{desk–work}\} \equiv E_\alpha\{\text{Sit–still}\} \land \{E_\varphi\{\text{Basal}\} \lor E_\varphi\{\text{Focused}\} \lor E_\varphi\{\text{Digestion}\}\}
\]

which would be read as “Desk-work is composed of the Sitting still activity, and of at least one of either Basal, Focused or Digestion physiology”.

The States, on the other hand, represent low-level hypotheses which are verified by comparing a model to the sensor data. As such, they possess knowledge as to which sensor input is needed, with which pre-processing, and computation models to evaluate the hypothesis’ plausibility (see Sect. 5.1). The States can also be given an expected duration (e.g. for Digestion).

Our system evolves over time, following the monitored person’s actions and physiology. Thus, most transitions, from a situation to another, are detected first by a change in the sensor data, and therefore, in the differences between the State hypotheses’ models and the input data. This led us to choose to handle these transitions at the State level, by introducing, for each State hypothesis, a set of successors: other States which can follow after a change in observations. For each State, this set is ordered by a distance \( d \), defined as a measure of how different two State data models are (e.g. Drowsiness is closer to Basal than to Exertion), so as to generate only the most relevant new hypotheses when a change is detected (see Sect. 4).

These States and Micro-scenarios thus form a situation network, as the example of Fig. 2 shows, with two kinds of edges: \( \mu S \) composition (in green, dashed) and State succession (in red, solid).

![Fig. 2: Example of small situation network.](image)

### 3.3 Alarming Situations

At this stage of our studies, we consider four kinds of alarms; this is preliminary work, but it gives a basis upon which to later build. As stated before, our purpose
is not to derive a diagnosis for a patient, but to provide estimations regarding the user’s condition. In this context, an alarm is approached as a multi-faceted notion, related to a situation rather than a single physiological model:

- A **Value Alert** can occur when “human thresholds” are passed (e.g. \( HR = 0 \)). As this is a simple filter, we do not concern ourselves with it as it can be handled by the sensors.
- \( \mu S \) **Alerts** happen when Micro-scenarios known to be alarming are verified: a simple example would be that of a person lying face-down with very high heart and breath rates. This is handled by \( \mu S \) agents (see Sect. 5).
- For **Scenario Alert**, we wish to measure a semantic distance between the current hypotheses and the planned Scenario, which would be a task for a probe (see Sect. 5.4).
- **System Alert**: the system is unable to find any likely hypotheses.
- **Hardware Alert**: the detection of sensor failure.

An ambulatory health monitoring system cannot, at any given time, decide on a single likely situation. Therefore, we need to be able to handle multiple hypotheses, on different time scales (to monitor both immediate and long-term hypotheses), with different sensor input.

### 4 Multiple Hypotheses Management

![Focus, Anticipation, Exploration](image)

Faced with a potentially large number of hypotheses, we need to select a relatively small number of hypotheses to evaluate and compare. We introduce a **confidence value** to evaluate each hypothesis, and the following mechanisms to navigate between hypotheses, on each level of abstraction, as shown on Fig. 3:

1. **Focus** is a higher level of abstraction suggesting something at a lower level: the Scenario steers the creation of \( \mu S \) hypotheses, which need specific States to be verified; these States then require specific data (see Sect. 5.2). This is a top-down method.
2. **Anticipation** occurs when a State agent’s confidence value drops, reflecting a change in the data it analyses. This drop is used to choose which new State hypotheses to generate, among its successors using \( d \) (as defined in Sect. 3.2), and the aperture \( A \) described in Sect. 5.2.
3. *Exploration* represents the need to link a State to a meaning in the working environment: States with a high confidence value generate Micro-scenarios based on prior knowledge such as the example of Fig. 2, if none of their possible µS already exist. This is purely bottom-up, and allows to widen the hypothesis set.

We therefore have two different prediction approaches, where *Anticipation* handles changes in the data while *Exploration* is based on opening the hypotheses to wider possibilities. The example shown in Fig. 3 is to be read as follows:

1. **T0**: the Scenario focuses the system by generating µS_{desk−work}, which likewise generates Eϕ_{basal} and Eα_{sit−still} to read their confidence values; these States require specific data and so choose the sensors’ focus.
2. **T1**: a change in the physiological data occurs and Eϕ_{basal}’s confidence value drops; it anticipates its possible replacement, in this example, by Eϕ_{eating}.
3. **T2**: after a verification period δexplo, Eϕ_{eating} considers itself very plausible and reflects this by exploring a broader meaning with regards to the working context, through µS_{meal} and µS_{tea−time break}. The Meal Micro-scenario can then be compared to the Scenario’s expectation of a “Lunch” to happen at some point.

5 Multi-Agent System

Each hypothesis at either µS or State level is represented by an agent (which we call Hypothesis Agents, or HA), whose role is first to compute a confidence value, and to choose a course of action from its evaluation. Agents are useful here because the hypotheses are to be evaluated with regards to specific data (either confidence values or observations from the sensors), thus giving them a confidence value that is computed autonomously. This “absolute” confidence value is then compared with thresholds that are dynamically set according to the global population of hypotheses. In other words, evaluating a given hypothesis (for example, that a person is digesting) does not impact the evaluation of another (for example, that this person is sitting, or speaking). A multi-agent system is well-suited to our need of autonomously-evaluated entities, with an adaptive regulation of the population as a whole.

One of the givens of our problem is that the State models can be heterogeneous; for example:

a. “digestion” is defined by a duration, and the contextual condition that it can only happen after “eating”.

b. The physiological state of “exertion” is characterized by minimum expected values for heart rate and breathing frequency, while “ingestion” is recognized by mean values of these signals, and a high variability on certain data.

To handle these models’ heterogeneity, we add two types of agents to the common HA architecture: Probes (see Sect. 5.4), and Data factories, which are
agents coupled with each sensor. They possess a library of transformation algorithms to both denoise the data, and process them according to each State agent’s current, specific needs (for example, some agents need an instant value, while others may require a standard deviation) on a per-demand basis. They are not further described here.

To differentiate between hypotheses, we introduce a confidence value $c$ which takes values in the same $I_c \subset \mathbb{R}$ for each HA (so as to be compared to one another); the closer the data are from the hypothesis’ model, the higher is $c$.

### 5.1 Confidence Value Computation Models

We built the system with the assumption that States’ data models would be known, and that the methods for computing their confidence values from the observed data could be quite heterogeneous, although comparable. The Micro-scenario models are, on the other hand, combinations of States. Simply put, our architecture is built on the basis of:

- **State** hypotheses, with confidence values $c_E$ computed from the observed data through data fusion, each with their own model thereof.
- **Micro-scenario** hypotheses, with confidence values $c_{\mu S}$ computed from its composing States’ $c_E$ values.

In our current system, to compute State confidence values, we extract a set of statistical features (mean, standard deviation, median, ...) from each type of observed signal. For signal $s$ (heart rate, breath frequency,...), the $j$th feature is denoted $f_s^{(j)}$. Then, we define confidence value as the average value of distances (noted $\rho$) between evaluated features and those provided by states models $m_{E_i}$, denoted $f_s^{m_{E_i}}$. $S$ is the number of observed signals and $J_s$ is the set of features extracted from signal $s$:

$$c_{E_i} = \frac{1}{S} \sum_{s=1}^{S} \sum_{j=1}^{J_s} \rho(f_s^{(j)}, f_s^{m_{E_i}}).$$

To evaluate Micro-scenario confidence values based on their State compositions, we assume that States are independent and that we can interpret their (normalised) confidence values as their existence probabilities, so that we can apply probability rules:

$$c_{\mu S} = c_{\mathcal{E}_1} \times c_{\mathcal{E}_2} \quad \text{if} \quad \mu S = \mathcal{E}_1 \land \mathcal{E}_2,$$

$$c_{\mathcal{E}_1} = c_{E_i} + c_{E_j} - c_{E_i} \times c_{E_j} \quad \text{if} \quad \mathcal{E}_1 = E_i \lor E_j.$$

This approach relies on the knowledge of models for each State (see Sect. 3.2), to be able to compare features extracted from the observed data. The models are assumed to be known and tailored to a subject’s physiology. This is a preliminary work and we are currently working on establishing a Bayesian approach exposed in [AF11], to learn models in an off-line process and compute the confidence values in an on-line process.
5.2 Available Information and Data

Table 1: Available Information.

<table>
<thead>
<tr>
<th>Current Loggs</th>
<th>Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current hypotheses: for each agent:</td>
<td>Global parameters: ( n_t, A, {\delta}_{\text{actions}} ), {thresholds}</td>
</tr>
</tbody>
</table>

Table 1 shows the three domains of available information for the system: the current system status, in a Blackboard through which the Hypothesis Agents (HA) share their confidence values; the Logs, containing traces of each agent action and past confidence values; the input sensor data used by the State agents. Here, \( n_t \) is the number of current HA at time \( t \), and the other parameters are used for the adaptive control of the HA population.

- \( n_t \) is the current number of agents
- \textit{aperture} \( A \): an indicator ruling how much leeway a State agent has when generating new hypotheses by anticipation and exploration.
- \( \{\delta\}_{\text{actions}} \) : the \( \delta \) durations represent the time an agent needs to ascertain its Confidence Range \( C_r \) (see Sect. 5.3) before undertaking an action.
- The \{thresholds\} are dynamically adjusted, for \( \mu S \) and State HA’s, and determine the limits of the Confidence Ranges.

5.3 Hypothesis Agents

The very purpose of evaluating hypotheses is to decide which among those are plausible. We introduce a Confidence Range noted \( C_r \), used by each agent to decide on its own course of action. It is an indicator of which hypotheses are considered likely, and which will be discarded, as can be seen in Sect. 6. \( C_r \) can be either \textit{High}, \textit{Medium}, or \textit{Low}, with two thresholds (high and low), which are dynamically adjusted, according to the current agent population, as a means of adaptive control (see Sect. 5.4). Figure 4 shows that an agent’s life cycle consists in periodically computing its confidence value from some input data, thus determining \( C_r \), which in turn determines action with regards to time spent in this confidence range (the \( \delta \) values, as seen in Sect. 4). The \textit{State} have specific action capabilities, through the \textit{Exploration} and \textit{Anticipation} mechanisms, shown here on Fig. 4 and described in Sect. 4).

Each agent has a model to compute its confidence value, and parameters to rule its actions:

\[
Agent = \{Name, Model, c, C_r, \{input\}, \{P\}, decision, f_{op}\}
\]
– $c$ is the confidence value ($c_E / c_{\mu S}$).
– $C_r$ is the confidence range.
– the input is, for the States, the data from the sensors’ data factories, and for the $\mu S$, the confidence values of its composing States.
– $\{P\}$ are the control parameters: thresholds, $\delta$ durations, and $A$.
– the decision is shown in Fig. 4 and in Sect. 4: various actions undertaken according to $C_r$.
– $f_{op}$ is the agent’s operating frequency.

### 5.4 Probes: Adaptive Control and Specific Tasks

<table>
<thead>
<tr>
<th>Ambient Environment</th>
<th>Prior Knowledge</th>
<th>Agent Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature, sound, weather.</td>
<td>Scenario, “mission”.</td>
<td>present and past system output (Blackboard and Logs): agent creations, confidence values...</td>
</tr>
</tbody>
</table>

We denote as probes, agents which operate independently from the Hypothesis Agents (HA). Table 2 shows the probes’ role: providing a contextual frame for the dynamic population of autonomous agents. These probes act either constantly, providing information about the global system status, or on a per-demand basis, to handle specific non-recurring requests made by the HA. Thus, we define three kinds of probes.

![Evolution of an agent’s confidence value.](image)

A Regulation Probe controls $A$, and sets the $\delta$ durations and $C_r$ thresholds for States and $\mu S$ to regulate the agent population. Its decision is based on information such as $n_t$, or their mean confidence value. For example, if there are
too many agents at a given time, its role will be to limit hypothesis generation:

Fig. 5 shows an example of an agent’s confidence values over time, where exploration happens at times $t_1$ and $t_2$; at time $t_3$, the agent does stop die because the probe changed the low threshold, so that the agent goes back up to the medium confidence range between $t_4$ and $t_5$. At $t_5$, the agent’s continued existence (and the subsequent agent creations at $t_7$) are a direct result of this probe’s decision.

The Consistency Probe searches through the logs for a specific occurrence (for example, upon creation, the Digestion agent requests that this probe check whether Eating recently had a high confidence value, and stops otherwise). Searching through the logs is a link between the HA and their context which can be computationally heavy, and is performed therefore only when needed.

An Alarm Probe is planned, to measure a semantic distance between the planned Scenario and the Micro-scenarios which have, at each time $t$, a high confidence value: it is still ongoing work, with the aim of detecting potentially alarming deviations from the expected plan.

6 Experimentation and Preliminary Results

Six healthy volunteers participated in the study, providing informed consent. It was approved by the CHU Grenoble’s ethics committee. The data from only one user were used in this section. Heart and respiration rates, body temperature, signals from 3D-accelerometer were recorded with a 5s period. The subjects were asked to come with work and lunch between 11 a.m and 1 p.m and given a Scenario of “daily life” steps. The data were used as shown in Fig. 6 (ground truth underlined in black, from the experimenter’s annotations), with a restricted situation network of 16 States and 9 µS and personalized data models. This illustrative example was implemented using java threads.

The first observation to be made is that the system correctly identifies the ground truth in this example, with very little ambiguity.

The hypothesis generation mechanism is based on both transitions and loose theoretical guidance. In both cases, the creator agent must often make tries not once but on a certain period of time. This may lead to a rather large number of agent creations, with often low confidence values. These new hypotheses are quickly discarded (in blue on Fig. 6: they are maintained only for a maximum initialization period which, for each hypothesis, is the minimum duration needed to evaluate this hypothesis). These creations can result in changes in the dynamic thresholds (as in the example of Fig. 5).

We observe that Micro-scenarios such as coffee-break, which are composed of large sets $E_i$ of possible States (see (1) in Sect. 3.2), representing transitory, chaotic combinations of States (here, coffee-break models a person drinking, talking, moving around or staying put, in no particular order), are not, with our current model, easily distinguishable from other Micro-scenarios composed of some common States. This results in some ambiguity in Fig. 6, which was ex-
Fig. 6: Agents’ confidence ranges over time, for one person.

expected as a premise; this is mitigated by the semantic closeness of the ambiguous Micro-Scenarios. Further work will aim at defining a measure of semantic distance between Micro-scenarios, in order to both compare them to the expected Scenarios, and to be able to regroup numerous and similar Micro-scenarios into a more generic one if need be.

Having made these first observations, we will now have to greatly increase our situation network, in order to demonstrate the efficiency we aim at, concerning the management of multiple hypotheses, and considering constraints such as real-time decision-making and power efficiency.

7 Discussion and Perspectives

In this paper, we have laid the theoretical foundations for a multi-agent, multi-hypotheses based personalized health monitoring system. With encouraging preliminary results, we will now aim at integrating a Bayesian-based confidence computation method, so as to be able to compare and choose the most effective approach. Hidden Markov Models are a robust approach, and can provide a common basis for the models (including sound machine learning processes), while our current heterogeneous States are more flexible and allow for a wide range of possibilities, but are more ad-hoc and based on human expert knowledge.

We wish to stress here the role that prior knowledge has in the current state of our work, both at the level of State data model, and of the situation network. Errors in either would greatly reduce the system’s effectiveness. However, the idea is to be able to manage a rather large number of possibles; therefore, further work will imply enlarging the situation network: we will then be able to effectively compare the computational complexity of this approach with other, Bayesian-based techniques, which have a risk of combinatorial explosion (it should however
be noted that the real-time aspects are not necessarily an issue, since the data from the sensor suite come at intervals of five to fifteen seconds), as well as combinations of both: the very purpose of our exploration mechanisms is to reduce the number of currently studied hypotheses to a manageable level.

With a larger situation network, the need for consistency checking will greatly increase, especially in order to limit the computational costs. The use of probes to reify contextual constraints on the agent population (e.g. such as filters [BB10]), will be developed further.

References