A Multi-Hypothesis Monitoring Architecture: Application to Ambulatory Physiology

Benoît Vettier, Laure Amate, Catherine Garbay

A Multi-Hypothesis Monitoring Architecture: Application to Ambulatory Physiology

Benoît VETTIER [a], Laure AMATE [b] and Catherine GARBAY [a]
[a] *Laboratoire d'Informatique de Grenoble, UJF-CNRS, France*
[b] *ISEN, Brest, France*

Abstract. In this paper, we propose a normative Multi-Agent System to handle uncertainty in a monitoring application. It is based on the assertion that no single most-likely situation should be considered, thus requiring the management of multiple concurrent hypotheses. A decision is then made by comparing these hypothesized situations to requirements and expectations, thus detecting potential problems. This system uses a large knowledge base of interconnected situation models on several levels of abstraction. It is centered around the need to constantly reconsider which hypotheses should be evaluated, with regards to both the current data from the sensors and wider requirements in terms of efficiency and specific focus from an expected scenario. We propose both a generic concept, and a more specific system for human health monitoring, using ambulatory physiological sensors.

Keywords. Multi-agent systems, Knowledge-based Information Systems, AI and Medicine, Hypothetical Reasoning, Hybrid Normative Systems

Introduction

Monitoring human activity implies addressing two main issues: taking into account the influence of *context*, and the paucity of input data, which can take the form of noise, ambiguity, or model variability. We consider this *context* as a combination of the environment (physical level), the temporal succession of situations (data level), and a set of requirements and expectations (interpretation level). This paper describes a Multi-Agent System which aims at handling these issues, both in a generic manner and in the specific case of monitoring a person's health using ambulatory, physiological sensors.

It is based on the *Superco* project of the French *DGA* (Defense Procurement Agency), which consists in a soldier wearing a monitoring device on the battlefield, to provide real-time, personalized information about his or her health and mission-readiness to the commanders. This means that the available data are limited to what light, wearable, unobtrusive sensors can provide. These data are then processed and interpreted in various ways, to fit the commander's informative needs: here, generating alerts when a potentially-worrying situation is detected.

The paradigm of Multi-Agent Systems allows for multiple, heterogeneous entities to be handled through a unified communication/cooperation frame. The *heterogeneity* of agents is considered as a requirement, to encompass the variety of situations a person can find herself in; a large knowledge base of interconnected situation models can thus be explored, as a dynamic population of multiple hypotheses on several levels of abstraction. These hypotheses are compared to the successive sets of sensor data over time, with transition mechanisms to handle the need to change the currently-studied hypotheses when the input data changes. They are then interpreted, to decide whether to raise alarms or not. Hypothesis management, as a whole, is performed within a frame of high-level requirements regarding efficiency, efficacy and adequacy of the system behaviour. A normative multi-agent architecture [6] is proposed to model these requirements and adapt them dynamically as new situations arise.

It is considered as a given that a large number of situation hypotheses must not be evaluated at all times: because of computational cost, and more importantly, to ensure the usability of results: considering ambiguous data (and models), too large a number of hypotheses would be considered likely, thus reducing the quantity of information each one holds. The hypotheses must be chosen and evaluated in light of a scenario, a path of successively-studied hypotheses, much like a word must be understood as part of a sentence and holds no assertive meaning on its own.

The context is prone to evolve, in terms of data and requirements: sensemaking is dynamically built as a match between captured data and elaborated expectations, in a framework of evolving requirements [20]. Adaptive control is therefore needed, which is the strength of a normative system such as this paper proposes: on one hand, the high-level (end-user) requirements can be dynamically updated, and on the other hand, operational rules adapt the parameters to the changing needs of the sensor data.

In light of this, we discuss, in this paper, some experimental results in the recognition of Activities of Daily Life (ADL), with the purpose of examplifying the proposed system's expressiveness, dynamics and regulation potential. The goal is then to detect discrepancies between hypothesized situation(s) and expectations, rather than trying to detect a single most-likely situation.

1. State of the Art

Monitoring Human Activity is known to be a complex task, whichever sensors, environments or goals are being considered. This is due to constraints on the available data, which are both noisy and ambiguous [21], but also to the fact that some crucial elements are *not quantifiable or measurable via sensors* [19]. In addition, the recognition process must draw on some a-priori knowledge about the objects at hand, since the way an object is "looked at" depends on what this object is expected to be [9]. Thus must all monitoring applications take the widely-studied *context* into account. This context is often hypothesized rather than sensed, which means that the interpreted situations carry more information than the sensors can account for. Such loosely-focused has a very wide spectrum of possible scenarios to choose from, which implies choices about which situations to consider (*what* to try and perceive).

In return, evaluating a situation model can result in *feedback* about the context. Activity recognition must therefore be seen as a constant loop of perception and choices about *how* to perceive.

Reasoning with multiple, concurrent hypotheses is a way to reflect the data's uncertainty, and the weight of the unobservable context. These hypotheses are based on heterogeneous data models which represent both stable situations and transitory periods. These can either be learned offline (as a-priori knowledge), as in Model-Based detection [8], or built online as exploratory learning [9]. Data-driven methods are used for the latter, often based on Hidden Markov Models [1], but with a requirement for human expert annotation [11]. In the specific case of human health, robust and proven models must be used from the start, as no expert annotation would be available on-line. Learning these models is thus out of this paper's scope: the knowledge base is considered known, and large (but cannot be exhaustive [17], as Human Activity is a very wide-scoped concept). The goal is then to navigate this knowledge base, to choose which hypotheses to evaluate.

A common comparison framework is necessary to handle these multiple, heterogeneous and autonomous hypotheses. Context Spaces are used in [18], in which degrees of support for hypotheses are computed via abductive inference). Confidence values are computed in [22], with associated cross entropy for overall correctness. The resulting confidence value reflects the amount of contradiction, if any, between model and data.

Monitoring human activity does not boil down to considering constraints from the realm of data. As a decision-making process, it has to meet the expectations of the social body in charge of this monitoring. Through analytical modeling [10], these requirements must be considered in addition to notions such as *effectiveness*, *efficiency* and *adequacy of results* (which is linked to *entropy*: the quantity of usable information). This calls for a hybrid system mixing bio-physical laws of the human body with the human organization's frames. Decision frames in such systems are to be modelled at several levels [25]: *local* frames suited to the entities' routine, and larger, *global* frames to ensure the compliance of the system as a whole, to institutional norms as functional requirements and goals [5]. This implies the decomposition of the global problem into smaller, more manageable phenomena [2], inside a contextual frame which provides, on higher levels of abstraction, disambiguation for the levels closest to the strongly-focused data models. [24], [23] and [7] propose such decompositions of human activity.

Normative MAS are a class of Multi-Agent Systems in which declarative, rule-based coordination components (called Filters [4]) allow the modelisation of various kinds of norms. These *filters* provide a dynamic, adaptive frame supporting complex control strategies, promoted for long by many authors [16] [12] [15]. They allow to separate the a-priori knowledge (the data models) and the algorithms (the enforcement of norms), thus complying with the guidelines to designing a monitoring platform defined in [3].

This makes for a complex decision paradigm, where the goal is no longer to recognize specific events but rather to detect discrepancies between observations and expectations: sensemaking is not a state of knowledge [13], but rather a process of fitting data into a frame that is continuously replaced and adapted to fit the data. The proposed architecture is a way to cope with these issues.

2. Generic Multi-Hypotheses Monitoring

As stated before, we consider a large knowledge base of interconnected situation models. These situation models are built on several levels of abstraction, so that a global scenario can be decomposed into meaningful, complex steps, which are composed of simpler, data-driven models. To be able to navigate between hypotheses, we need to handle transitions: when the situation changes, so do the sensor data, and the system must therefore "resample" to adjust the currently-studied hypothesis population to the new data. All hypotheses are evaluated through a *confidence* value, which reflects its likelihood with regards to the current data and context.

The *situation network* is thus composed of all situation hypotheses, with both *horizontal* edges linking hypotheses of the same kind, and *vertical* edges for combinations on different levels of abstraction. Figure 1 shows an example with several hypotheses evaluated over time, with changing likelihoods. *Desk-work* is seen as a combination of *Sitting* and a *Basal* physiology; the latter sees its confidence drop (time t_1), and generates another one (*Ingestion*) to replace it.

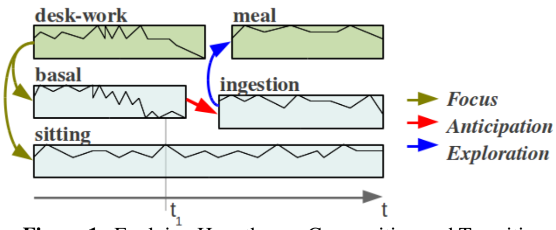

Figure 1. Evolving Hypotheses: Composition and Transition.

The mechanisms of *Focus* (evaluating data-level hypotheses according to a given context), *Anticipation* (echoing a change in the data by generating "successor" hypotheses) and *Exploration* (reaching upwards from data-level hypotheses to "high-level", meaningful hypotheses) are built into the Multi-Agent System (MAS) and use the links of the known situation network. These mechanisms must be regulated, so that the system generates an adequate number of relevant hypotheses: in addition to the situation network, we therefore need knowledge and rules to ensure that the hypothesis generation is open enough, while still retaining a high informative level (that is, avoiding to drown the system in a flood of inseparably likely hypotheses). Such *operational* rules are embodied by Filters, as described in Section 4.

Moreover, should the hypothesized situation stray from an acceptable or, at least, expected frame, rules must be in place for the system to react and generate some alert. This is the application of high-level expectations and requirements; these can also be written in the form of rules, so that the MAS uses a unified Filter engine to apply both operational (regulation) and *institutional* (expectations) rules.

These rules can be very varied in kind, and can be added, removed or updated on-the-fly, either through autonomous adaptive control, or by human intervention.

3. Multi-Agent System

The multi-hypothesis management architecture is built upon a dynamic population of hypothesis-agents (called $\mathcal{H}$ in the rest of the paper), which can be referred to as either *agents* or *hypotheses* with the same meaning: the combination of a data model, meaning, algorithms, and current output values.

The hypothesis-agents's operating cycle (see Fig. 2) consists in waiting for notifications to perform two kinds of actions: evaluating the likelihood of the hypothesis given the current data (*verification*), and choosing which new hypotheses (if any) should be evaluated next (*prediction* to replace the current hypothesis when found to be weak).

The agents share information through a Blackboard, and Filters [4] apply sets of rules to ensure the system's compliance to an adaptive, dynamic set of norms.

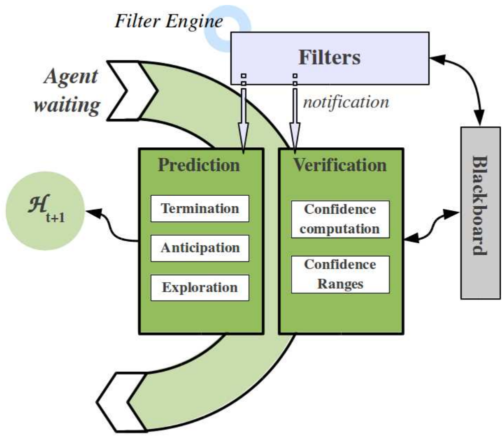

Figure 2. Hypothesis Agent Cycle.

3.1. Hypothesis Agent Definition

A hypothesis-agent is a tuple $\mathcal{H} = \{K, \chi, \nu, C, R, t_r\}$, with:

- K: the components (either sensor data or other agents' output info) from which the confidence value is computed
- χ: verification methods (to compute a *confidence value* from the *components*)
- ν: relations to other hypotheses (see *Prediction*)
- C: confidence value
- R: confidence range (low, $medium$, $high$)
- t_r: durations spent in each R

3.2. Knowledge, Data and Information

Considering [21]'s nomenclature, we consider the following:

- *Data*: coming from the sensors; used by low-level agents to verify basic hypotheses. The data generation is independent from the system's operation but different kinds of pre-processing can be applied when needed.
- *Knowledge*: a situation network (oriented graph) of hypothesis models, from data-level to contextually meaningful situations. The Knowledge part is "read-only".
- *Information*: all of the agents' output, particularly confidence values (resulting from data abstraction).

Figure 3 shows how the architecture is built around the Blackboard, where Information and Data are shared. The hypothesis-agents are the seats of all comparison between Information/Data and Knowledge. Both Information and Data propagation are subject to rules applied by a kind of Filter (see Section 4), which aims at reducing unnecessary calculus and results in a forced synchronization of input data. The final goal is the generation of alerts.

3.3. Verification

A verification step is needed when the sensors provide new data: this means re-evaluating the confidence of each current hypothesis linked to these data. The goal here is not so much to decide which hypothesis is most likely, as to determine a set of "likely enough" hypotheses which can be compared to global expectations.

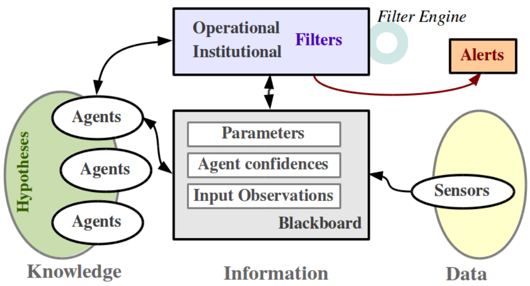

Figure 3. Multi-Agent Architecture.

To distinguish between the plausible hypotheses, those that are possible and those that are completely off, three confidence ranges are defined (high, medium and low). Predictive actions will be decided upon depending on the confidence range an agent is found to be in; these ranges are separated by thresholds T_{low} and T_{high}, which can be dynamically adapted by an operational filter (see Section 4):

$$\mathcal{H}.R = \begin{cases} low \;\; if \;\; \mathcal{H}.C < T_{low}, \\ high \;\; if \;\; \mathcal{H}.C > T_{high}, \\ \quad else \;\; medium \end{cases}$$

3.4. Prediction

Over time, the hypotheses' plausibility (confidence) varies as new sensor data arrive. Should a hypothesis be proven unlikely by the new data, it must be replaced by a new set of relevant hypotheses. Choosing these new hypotheses requires taking into account both some a-priori knowledge of which situations $\{\mathcal{H}\}_{t+1}$ can follow $\mathcal{H}_t$ at time t, and the difference between observations Y_{t-1} and Y_t. This can be done by computing a *distance* d_{ij} on the *horizontal* links of the situation network (see Section 2), between $\mathcal{H}_i.\chi$ and $\mathcal{H}_j.\chi$. This is called *Anticipation*, and happens when an agent is found to be in the R_{low} or R_{medium} ranges for a given duration.

As these new *anticipated* hypotheses $\{\mathcal{H}\}_{t+1}$ are the result of variations in the sensor data, it takes place mainly in the lower-level hypothesis-agents, thus generating new data-level agents which need to be anchored into a wider contextual meaning: higher-level hypotheses. This mechanism, called *Exploration*, is illustrated in Fig. 1, where the data-level *Ingestion* hypothesis creates a wider *Meal* hypothesis. *Exploring* agents create other agents they are components of: $\mathcal{H}_{exploring} \in \mathcal{H}_{explored}.K$.

It must be emphasized that these methods rely on a rich situation network, which includes not only a large number of data models, but also links between hypotheses: compositions between abstraction levels, possible transitions, and distances between models (in terms of both expected data and semantics). This knowledge base must be learnt and computed off-line (which is out of this paper's scope).

3.5. Hypothesis Patterns

The successive Prediction steps can be seen as the creation of hypothesis paths, or timelines. These simultaneously-evaluated concurrent timelines can be compared to an expected scenario: either to recognize specific patterns, or to detect that the situation strays from the expectations. As previously stated, such comparisons require a measure of distance between hypotheses; only high-level hypotheses (full, complex situations rich in contextual meaning) should be built into timelines, as expectations regarding a planned activity are expressed in terms of meaningful situations rather than data-level models.

Moreover, a given hypothesis can be reinforced (its confidence value increased) if it is the last element in a path of strong past hypotheses, meaning that the timeline it results from is more likely than a string of weak, loose hypotheses.

4. Normative MAS

A Normative MAS is a society of autonomous agents, with organizational rules governing these agents' activity and what information is available to them. The filters, as defined in [4], are a set a rules, defining requirements and activation contexts. An autonomous filter engine, with its own operating frequency, applies these rules, which result in messages (activation notifications) sent to the agents. These rules can be given frequencies, orders of precedence, and priorities. These filter parameters, and the rules themselves, can be dynamically modified so as to shift the system's focus. This can be done either through autonomous adaptive control, or by a human intervention (for example if a supervisor decides to change the global requirements). Moreover, the heterogeneous agents can use different sets of filters, as the agents are responsible for their subscription to relevant filters.

Table 1 gives some filter rules examples, which are written with conditions, messages (*what to do*) and targets (the agents the messages are sent to). n_{red} is the number of unlikely hypotheses (confidence ranges R_{low} and R_{medium}, as opposed to R_{high}, the likely hypotheses), while N_{ef} is a parameter of the *Efficiency* filter.

4.1. Parameters

The notion of transition from a situation to another is closely linked with the notion of duration: generating a new hypothesis is merely a loose assumption of a data change's relevance at a given time: therefore, hypotheses are given *inertia* durations before the predictive actions (anticipation, exploration, termination) are triggered. This implies introducing parameters which we call δ (and which are compared to the agents' t_r timers). Modifying these parameters will adjust the system's *aperture*; that is, generating enough new hypotheses to reflect the person's evolution, while minimizing the sensitivity to noise and allowing for the minimum period during which a hypothesis must be evaluated, for this evaluation to be meaningful.

As seen on Figure 3, one kind of filters is the *Operational* Filters, which control each hypothesis agent's actions, and maintain their consistency with the data. These are:

- *Verification* filters are tasked with propagating new input data when needed (rules are applied to check whether the data changed enough to compute new confidence values, and noise is filtered).
- *Prediction* filters apply the three predictive actions (defined in the previous section) by comparing the agent's confidence ranges and associated timers to the current thresholds.
- *History* filters are used to select which pieces of information deserve to be stored in the logs (so as to avoid drowning the useful traces), and browse the logs.

Filter	Target	Message	Condition
Op: Termination	all $\mathcal{H}_i$	*term*	$\mathcal{H}_i.R$ = *low* & $t_{low_i} = \delta_{term}$
Op: Anticipation	all $\mathcal{H}_i$	*ant*	$\mathcal{H}_i.R$ = *med* & $t_{med_i} = \delta_{ant}$
Inst: Efficiency	T_{low}	inc(T_{low})	$n_{red} > N_{ef}$
Inst: Efficiency	T_{low}	dec(T_{low})	$n_{red} < N_{ef}$

Table 1. Filter Rules: Examples

4.2. High-Level Requirements

The *Institutional* filters apply the global requirements by modifying the parameters according to a set of regulation rules:

- *Regulation* filters are dedicated to the system's adaptive control, by dynamically adjusting parameters. For example, the *Efficiency* rule may raise the T_{low} threshold to discriminate more against unlikely hypotheses, when these are too numerous (more than a N_{ef} number). Table 1 illustrates this filter's rules.
- *Alert* filters detect discrepancies between acceptable expectations (norms) and current hypotheses, to generate alerts as defined previously. Among these rules are, for example a semantic distance to the Scenario, or checking for unacceptable Micro-scenarios.

While the Operational filters provide a local, data-driven frame to handle the agents' routine operation, the Institutional filters ensure the system's regulation as a whole and fulfill its goal of detecting potentially unacceptable situations.

5. The Specific Case of Human Physiology

5.1. Data and Hypothesis Models

The *Superco* project is built around the idea of generating alerts when a person's situation is detected as deviating from acceptable expectations. This detection must use ambulatory, physiological sensors. In this case, Heart and Breath Rate, along with Skin Temperature and a 3-axis Accelerometer. Early results have shown that the data models must not only be tailored to each person (depending on fitness, gender, age...), but also adapted to several possible contexts (for example, the basal Heart Rate for a given person varies between morning and evening). This results in a great uncertainty and ambiguity, as different hypothetical situations could result in the same observations from the sensors in different settings. This Knowledge Base is considered known and does not change.

Moreover, a single situation can be characterized differently depending on a person's habits (for example, the inability to stand still without growing restless), mood, state of hunger or fatigue... Therefore, both the hypotheses' data models, and the links between hypotheses, must be adapted to each person's specific physiology: there are no generic models for physiological values such as Heart Rate, except for naive viable ranges.

For this application, we define two levels of hypotheses: States, and Micro-scenarios, which are full, complex situations, seen as steps in a global scenario. A *Scenario* can thus be built as a succession of Micro-scenarios. For example, a *Meal* or a *Phone Call* could be parts of a Daily Life scenario. A Micro-scenario is a combination of broad meaning and a set of components which embody the simpler hypotheses composing the situation: Figure 4 shows an example of a Micro-scenario (*desk-work*) which is composed of two States (*basal* physiology and *sitting* activity).

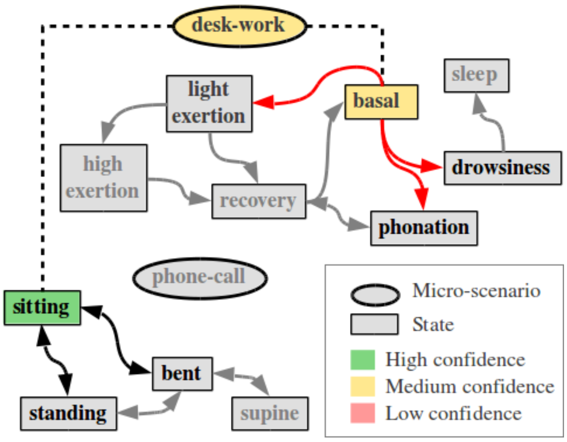

Figure 4. Situation Network: Simple Example for ADLs.

While Micro-scenarios are context-driven, *institutional* models which represent complete, meaningful situations, States are data-driven, *operational* hypotheses regarding the sensor observations. Their components are the input data themselves, from which a confidence value is computed, using expected values in a given context. This confidence value can be separated into ranges, which define the predictive actions undertaken by the agents: in the example of Figure 4, the *basal* hypothesis is unlikely (*medium* confidence) and its agent will therefore create successors to replace it (*phonation*...).

5.2. Alarming Situations

The notion of *alert* dwells in the observer's eye rather than in the observed data itself: alerts depend on the context and the requirements. As such, their rules must be defined in a declarative and dynamic manner, for which we propose three main axes.

The *data-level* alerts are basic, universal and objective alerts: sensor failure, crippling levels of noise, and physiological values that are known as out of a person's viable bounds (such as a Skin Temperature below 25°C). These alert models are tailored to each person's specific physiology.

On the other hand, *scenario* alerts are full situations, taking the context into account, and which involve both common sense, and expectations. These alerts include situation hypotheses known as unacceptable (either a priori, or defined as such on-the-fly), or situations which simply differ from what the subject is expected to do (for example, a soldier on a given mission), without being a danger to his health by themselves. These *scenario* alerts are given various levels of importance, from a sudden, very alarming situation to a pattern of hypotheses which may be the early symptoms of a condition which would fully develop later.

Finally, *system* alerts are raised when the data interpretation fails to comply with requirements such as effectiveness, efficiency or entropy: the system's output would not be useful if too many hypotheses were simultaneously considered likely, for example.

6. Results and Discussion

This section presents some results obtained by simulating a real-time interpretation: a Java application reading, a posteriori, a set of input data which come from a monitoring belt providing synchronized Heart Rate, Breath Rate, Skin Temperature, and 3-axis accelerometry. Future data collection could include environmental data such as altitude, outside temperature, or sound volume, to enrich the interpretation of context. The data collection protocol was validated by an ethics committee. The data models were tailored to the subject, as per the knowledge constraints stated in Section 5. This was done in cooperation with a team of physiologists.

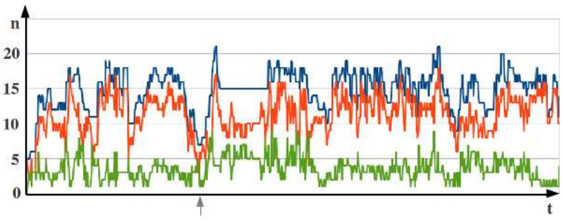

Figure 5. Number of Likely/Unlikely Hypotheses.

Figure 5 shows the number of hypothesis agents over a longer period of time (in blue: all, in red: unlikely $\mathcal{H}$, in green: likely $\mathcal{H}$), and illustrates the operation of the *Efficiency* filter (see Table 1): whenever the number of unlikely hypotheses reaches a given level, the parameters are modified so that the system becomes less tolerant, and this number subsequently drops. Point A on Figure 5 reveals such a drop, which is followed by a sudden increase. This highlights the system's dynamics: when the interpretation is unsatisfactory, new hypotheses are generated to replace the failing ones (both States and Micro-scenarios). Upon creation, the Micro-scenario *focus* (see Figure 1) will generate new State hypotheses as its *components*.

Figure 6. Influence of the Verification Frequency.

Tuning such a monitoring system implies a necessary compromise between sensitivity and computational cost (it is crucial to save batteries and match real-time constraints on a real ambulatory system). This balance may be found by off-line learning and by online adaptive control, particularly to focus the system's resources on periods of greater criticity [12]. For example, reducing the *Verification* frequency (increasing the Verification Filters' period, which can be done on-the-fly) reduces the system's tolerance to low-likelihood hypotheses: their confidences are smoothed towards R_{low}: Figure 6 shows a snapshot of two State hypotheses, with the same data, but with different f_{verif}. The green, yellow and red lines show, over time, the confidence ranges (high, medium, low). At time X, the *Phonation* hypothesis is destroyed (low confidence) only for the higher Verification period. It also mechanically results is fewer *anticipated* hypotheses (since there are less *anticipating* agents).

This highlights that tweaking one parameter has an effect on virtually every aspect of the interpretation mechanisms: a lower number of hypotheses means a different application of filter rules such as the *Efficiency* criterion shown on 1, which in turn has an impact on the tolerance to unlikely hypotheses.

Further work will focus on training the system's adaptive control rules. This may include the discovery of heuristics, as the parameters' interdependency may prove costly.

Moreover, the richer the interpretation system, the more parameters there are, and therefore the more complicated this interdependency becomes.

The numbers of likely and unlikely hypotheses (resp. green and red lines on Fig. 5) are indicators to different features of the system. The "green line" shows how well the system recognizes which situation the subject is in: if n_{green} is too low, it may mean that the system does not recognize the situation. But if n_{green} is too high, the interpretation is too fuzzy: no decision can be taken as to which situation is "true". However, if neither of these too-numerous likely hypotheses trigger any *alert*, then the interpretation still achieves the system's goal of verifying whether the person is in trouble. The green line can therefore be interpreted as relevant to the *Institutional* requirements.

On the other hand, the "red line" (unlikely $\mathcal{H}_i$) shows the system's operational activity. Too low a number of unlikely hypotheses would mean that the system is not *open* enough (the Prediction step is akin to re-sampling). Indeed, the unlikely hypotheses are evidence of the system's exploration of varied possible situations.

Depending on the situation's criticity, rules can be added or modified on-line, so that the system is more or less open. The driving idea here is to be able to focus the system's resources on the most informative/critical elements.

As to the number of hypotheses considered likely (n_{green}), the *False Positive* criterion (hypotheses considered likely but which are not true) is not entirely relevant (aside from the uncertainty previously mentionned): as it is the very nature of these data to be ambiguous, the goal is to detect the possibility of an alert.

On the other hand, a False Negative criterion can be applied to the generation of alerts. However, this is made difficult by the fact that the amount of possible alerts depends on the varying requirements (since the institutional filters can be dynamically modified by the end-user). Further work will focus on the specification of alert rules, especially with regards to a comparison between an expected scenario [14], and the various hypothesized timelines (as mentionned in Section 3.5).

7. Conclusion

We have proposed in this paper, a generic architecture to monitor human activity. Multiple interpretation hypotheses are processed concurrently, at several abstraction levels, by independent agents, based on data models and contextual knowledge (situations and scenarios at hand) that are acquired and learned off-line (using annotated data and human expert knowledge).

The system is designed to raise alerts that are grounded both in evidence from the data and expectations from the social bodies in charge of the monitoring. In addition, it has to behave in a way that is consistent with respect to operational requirements, regarding for example the range of hypotheses to be considered simultaneously. In this context, one major designing guideline is to provide declarative and separate modelling of these requirements. The second one is to provide dynamic adaptation capabilities, to ensure that hypothesis generation comply with current requirements, but also to ensure that these requirements stay up-to-date with regards to the evolving context: operational and institutional requirements are meant to evolve, depending *(i)* on the current number of hypotheses (and their confidence values), and *(ii)* on the possibility for alarms, or hints thereof. A normative multi-agent system is proposed in this perspective and some results discussed for ADLs. These results are preliminary, and more sophisticated models are needed. However, they highlight the system's expressiveness and adaptativity potential.

References

[1] L. Amate, F. Forbes, J. Fontecave, B. Vettier, and C. Garbay, 'Probabilistic model definition for physiological state monitoring.', in *IEEE Intl. Workshop on Statistical Signal Processing*, (2011).

[2] F. Amigoni, M. Dini, N. Gatti, and M. Somalvico, 'Anthropic agency: A multiagent system for physiological processes', *ARTIF. INTELL. MED*, **27**, 305–334, (2003).

[3] A. Avanzi, F. Bremond, C. Tornieri, and M. Thonnat, 'Design and assessment of an intelligent activity monitoring platform', *EURASIP J. Appl. Signal Process.*, (2005).

[4] F. Badeig, F. Balbo, and S. Pinson, 'A contextual environment approach for multi-agent-based simulation.', in *ICAART*, (2011).

[5] D. Baker, D. Bridges, R. Hunter, G. Johnson, J. Krupa, J. Murphy, and K. Sorenson, *Guidebook to Decision-Making Methods*, 2001.

[6] O. Boissier, F. Balbo, and F. Badeig, 'Controlling multi-party interaction within normative multi-agent organizations', in *Coordination, Organizations, Institutions, and Norms in Agent Systems VI*, (2011).

[7] P. Chahuara, A. Fleury, and F. Vacher, M.and Portet, 'Methodes svm et mln pour la reconnaissance automatique d'activites humaines dans les habitats perceptifs: tests et perspectives', in *RFIA*, (2012).

[8] MO. Cordier, P. Dague, F. Levy, J. Montmain, M. Staroswiecki, and L. Trave-Massuyes, 'Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspective', *IEEE transactions on systems, man and cybernetics*, **34**, 2163–2177, (2004).

[9] M. Felsberg, J. Wiklund, and Granlund G., 'Exploratory learning structures in artificial cognitive systems', *Image and Vision Computing*, **27**, 1671–1687, (2009).

[10] A. Freedy, M. Cohen, G. Weltman, and E. Freedy, 'A new tactical group decision analysis system (tgdas) combining analytical and naturalistic decision modeling', in *IEEE Conference on Collaborative Technologies and Systems*, (2007).

[11] T. Guyet, C. Garbay, and M. Dojat, 'Knowledge construction from time series data using a collaborative exploration system', *J. of Biomedical Informatics*, (2007).

[12] Barbara Hayes-Roth, 'An architecture for adaptive intelligent systems', *Artif. Intell.*, **72**(1-2), 329–365, (1995).

[13] G. Klein, J. Phillips, E. Rall, and D. Peluso, 'A data/frame theory of sensemaking', in *Expertise Out of Context: 6th Intl Conf. Naturalistic Decision Making*, (2006).

[14] C. Lesire and C. Tessier, 'Particle petri nets for aircraft procedure monitoring under uncertainty', in *26th Intl. C. On Application and Theory of Petri Nets and Other Models of Concurrency*, (2005).

[15] L. Merghem, D. Gaiti, and G. Pujolle, 'On using multi-agent systems in end to end adaptive monitoring', in *MMNS*, (2003).

[16] J. Meyer and F. Mili, 'Self-adaptive selective sensor network querying', in *SASOW '08: Second IEEE Intl. Conf. on Self-Adaptive and Self-Organizing Systems Workshops*, (2008).

[17] D. Minnen, I. Essa, and T. Starner, 'Expectation grammars: Leveraging high-level expectation for activity recognition', in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2003).

[18] A. Padovitz, S. Loke, A. Zaslavsky, C. Bartolini, and Burg B., 'An approach to data fusion for context awareness.', in *Fifth International Conference on Modelling and Using Context*, (2005).

[19] A. Pantelopoulos and N. Bourbakis, 'Prognosis - a wearable health monitoring system for people at risk: Methodology and modeling', in *IEEE Trans. on Information Technology in Biomedicine 14-3*, (2010).

[20] Simone Santini, 'Image retrieval', *IEEE Intelligent Systems*, **17**, (2002).

[21] A.S. Silvent, M. Dojat, and C. Garbay, 'Multi-level temporal abstraction for medical scenario construction', *International Journal of Adaptive Control and Signal Processing*, **19**, 377–394, (2005).

[22] B. Thomson, K. Yu, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, and SJ Young, 'Evaluating semantic-level confidence scores with multiple hypotheses.', in *Interspeech*, (2008).

[23] J. Bigham U. Naeem, 'Activity recognition using a hierarchical framework', in *2nd Intl. Conf. on Pervasive Computing Technologies for Healthcare, Ambient Technologies for Diagnosing and Monitoring Chronic Patients Workshop*, (2008).

[24] B. Vettier, L. Amate, C. Garbay, J. Fontecave, and P. Baconnier, 'Managing multiple hypotheses with agents to handle incomplete and uncertain data', in *URMASSN International Workshop*, (2011).

[25] K. Weber and M. Glynn, 'Making sense with institutions: Context, thought and action in weick's theory', *Organization Studies*, **27**, (2006).