Reading Agendas Between the Lines, an exercise

Giovanni Sileno and **Alexander Boer** and **Tom van Engers**[1]

Abstract. This work presents elements for an alternative operationalization of monitoring and diagnosis of *multi-agent systems* (MAS). In contrast to traditional accounts of model-based diagnosis, and most proposals concerning non-compliance, our method does not consider any commitment towards the individual unit of agency. Identity is considered to be mostly an attribute to assign responsibility, and not as the only referent that may be source of intentionality. The proposed method requires as input a set of prototypical *agent-roles* known to be relevant for the domain, and an *observation*, i.e. evidence collected by a monitor agent. We elaborate on a concrete example concerning tax frauds in real-estate transactions.

INTRODUCTION

In previous works [2, 3], we have presented a *model-based diagnosis* view on complex social systems as the ones in which public administrations operate. The general framework is intended to support administrative organizations in improving responsiveness and adaptability, enabled by the streamlining of use cases and scenarios of non-compliance in the design cycle and in operations. This paper focuses in particular on the *operationalization* of model-based diagnosis (to be used in operations, and therefore supporting responsiveness) and differs from the previous papers in granularity, as it provides a specific example of implementation. Note that even if we apply the proposed method to identify the occurrence of non-compliance, it may be used in principle for any other pattern that may be of interest for the organization.

The paper is organized as follows. § 1 provides a general introduction to diagnosis, and to what we intend as diagnosis of social systems; § 2 presents an overview on the various literature in AI about model-based diagnosis; § 3 introduces the case study (sale transactions of real-estates), identifying prototypical scenarios of interest; § 4 concerns the actual exercise of operationalization of monitoring and diagnosis, providing insights and directions for future developments.

1 DIAGNOSIS OF SOCIAL SYSTEMS

In general, a diagnostic process is triggered if there is the presumption that a *failure* occurred in the system. However, what counts as a failure depends on the nature and function of system.

In case of a *designed artifact*, the system is generally associated to a set of requirements, and, at least at the moment of production, to an implementation model—a *blue-print*. A *failure* becomes manifest when there is an inconsistency between the form/behaviour that is observed and what is expected from that artifact. The failure may be at the *design level*, when the implementation does not meet the design requirements; or at the *operational level*, when one of the subcomponents has failed, and propagated its failure to the system.

In case of a *social system* (natural or artificial), the internal mechanisms of social participants are unknown and typically inaccessible. For instance, we are not able to fully know what is in the mind of a person, nor how someone's mind actually works (not even our own).[2] Nevertheless, we still *do* apply (when it is relevant to do so) a *theory of mind* to explain and interpret our own or others' behaviour, by referring to notions as beliefs, desires, and intentions. If we assume that the application of this stance is viable, then, when something goes wrong in a social system, i.e. when someone's expectations about the behaviour of someone else are not met, this means that something went wrong at as informational, motivational, or deliberative level of at least one individual.[3] In order to identify the wrong, however, we have to consider the requirements associated to the system. A first filter of discrimination could be obtained by referring to normative directives: prohibitions and obligations correspond respectively to negative and positive requirements. This would be sufficient, if the contextualization of a generic norm in an actual social setting was straightforward. However, as the existence of the legal system shows, this is far from being the case: the *qualification* of actions, conditions, people and the *applicability* of rules build up the core of the *matter of law* debated in courts. Thus, in an *operational setting*, rather than norms, we need to refer to adequate abstractions of cases, making explicit factors and their legal interpretation; in this way, we handle *contextualized normative models* that can be directly used to discriminate correct from faulty behaviour, all while maintaining a legal pluralistic view.[4]

1.1 Deconstructing identity

Current approaches of diagnosis on MAS consider social system components (software agents, robots, or persons) as individual intentional entities, i.e. following an assumption that could be described as "*one body, one mind*" (see references in § 2.1). In contrast, we assume that intentional entities may transcend the individual instances of the agents. In the case of a *combine* (e.g. in sport, when a player makes an agreement with a bidder on the results of a match) or similar schemes, the collective intentional entity that causes and explains the resulting behaviour is placed behind the observable identities.

[1] Leibniz Center for Law, University of Amsterdam, the Netherlands, corresponding author: g.sileno@uva.nl

[2] In the words of Chief Justice Brian (1478): "for the devil himself knows not the thought of man".

[3] This is true also in domains where the law imputes *strict liability*, i.e. where the claimant only need to prove the occurrence of the *tort*, and not of a *fault* (negligence, or unlawful intent) in the agent who performed the tort. In these cases, the law discourages reckless behaviour, pushing the potential defendant to take all possible precautions. In other words, in strict liability law ascribes fault *by default* to the agents making a tort.

[4] This may be useful for practical purposes: a public administration may for instance use dissent opinions of relevant cases to further strengthen its service implementations.

Such an interpretation of intentionality has relations with the notions of coordination, coalition formation, and distributed cognition.[5] In addition to this "*one mind, many bodies*" scenario, we allow that an agent may interleave actions derived by a certain strategy with actions generated for other intents, independents from the first: the "*one body, many minds*" case may apply as well.

1.2 Diagnosis as part of a dual process

Monitoring agents (e.g. tax administrations) are typically continuously invested with a stream of messages (e.g. property transfer declarations) autonomously generated by social participants. Clearly, they would encounter a cognitive overload if they attempted to reconstruct all "stories" behind such messages.

In affinity with Dual Process theories of reasoning, we may distinguish a *shallower*, less expensive but also less accurate mechanism to filter the incoming messages; and a *deeper*, more expensive, and accurate mechanism to analyze the filtered messages, possibly performing further investigative actions. The first, implemented as a *monitoring* task, is designed by settling what is interesting to be monitored, and which are the threshold conditions that identify *alarming* situations. The second, implemented as a *diagnostic* task, is triggered when such (potentially) alarming situation are recognized, and possibly starts specific courses of actions to look for other clues discriminating possible explanations (diagnostic and non-diagnostic). Note that the two tasks are intimately related: they are both constructed using expectations of how things should go, and of how things may go wrong. Furthermore, planning builds upon abilities, which can be reinterpreted as expectations of how things may go performing certain actions in certain conditions. From a practical reasoning point of view, *planning, monitoring and diagnosis are parts functional to a whole, and the practical reasoning of an agency cannot but be disfigured if one of these functions is neglected.* In other words, all effort that a public administration puts into simplifying the operations in the front-office of service provision (e.g. diminishing the evidential burden on the citizen) should be coupled with effort in the back-office in support of institutional maintenance.

1.3 Side effects

The choice of investigative actions requires some attention as well. In the case of physical systems, *measurements* do not necessarily involve a relevant modification of the studied system (at least at a macro-level), and criteria in deciding amongst alternative measuring methods generally concern costs on opportunities. In the case of a social system, this cannot be the only criterion. For instance, if the target component suspects being under observation, he may adopt an *adversarial* or a *diversionary behaviour* protecting him from intention recognition actions (cf. [28]); he may also drop the unlawful intent as a precaution. In this work, we overlook the planning problem for evidence-gathering tasks taking into account these derived behavioural patterns.

2 RELEVANT LITERATURE

Model-based diagnosis is a traditional branch of study of AI (see e.g. [21] for an overview); it has reached maturity in the 1990s, and it has been applied with success in many domains, reaching a production level of technology readiness (see e.g. [7]). In the following, we retrace the main directions of investigation, highlighting where relevant the specificities of our problem domain.

2.0.1 Consistency-based diagnosis

Early approaches in model-based diagnosis used explicit fault models to identify failure modes (see e.g. [13]), but these evolved towards diagnostic systems based on descriptions of correct behaviour only. Practical reasons explain this progress: in the case of electronic devices, manufacturers provide only descriptions of normal, correct behaviour of their components. Failure modes could be computed simply as inconsistencies with the nominal specifications (cf. [26] for a minimal set of faulty components, [14] for multiple faults configurations). This type of diagnosis is usually called *consistency-based diagnosis*. In short, by having models of correct behaviour of the system components and a topological model of their composition and knowing the initial state, we can predict the expected system state via simple deduction. If the observed output is different, we acknowledge a behavioural discrepancy, which triggers the diagnostic process aiming to identify the faulty components. Note that in this case, such components are deemed *faulty* simply because they do not behave according to their nominal specification: the 'negative' characterization is then constructed in duality to the 'positive' one (cf. *negation as failure*). In recent literature, these are also called *weak fault models* (WFM), see e.g. [35]. This approach entails important consequences: in consistency-based diagnosis, all fault models become equivalent, meaning that, from the diagnoser perspective, "a light bulb is equally likely to burn out as to become permanently lit (even if electrically disconnected)" [15].

2.0.2 Abductive diagnosis

Not surprisingly, the approach provided by consistency-based diagnosis is not fit for certain domains. In medicine, for instance, doctors do not study only the normal physiology of human organisms, but also how certain symptoms are associated to diseases; the hypotheses obtained through diagnosis are used particularly to *explain* given symptoms. In other words, 'negative' characterizations—*strong fault models* (SFM)—are asserted in addition to the 'positive' ones (cf. *strong negation*), rather than in duality to them. In the literature, in order to operationalize this approach, several authors have worked on explicitly characterizing the system with faulty models, starting a line of research which led to the definition of (model-based) *abductive diagnosis* (see e.g. [11], [8]).

2.0.3 Type of diagnosis per type of domain

We can sketch two explanations of why certain domains refer to consistency-based diagnosis, and others to the abductive diagnosis. The first explanation is built upon the use of negation. The former approach takes a *closed-world assumption* (CWA) towards the system domain, while the latter considers an *open-world assumption* (OWA), reflecting the strength of knowledge and of control that the diagnoser assumes having. Reasonably, engineering domains prefer the former (everything that does not work as expected is an error), while natural and humanistic domains usually refer to the latter (there may be a *justification* for why things didn't go as expected). The second explanation considers the different practical function for which diagnosis

[5] cf. [17]: "A central claim of the distributed cognition framework is that the proper unit of analysis for cognition should not be set *a priori*, but should be responsive to the nature of the phenomena under study."

is used in the domain. While by applying consistency-based diagnosis we can identify (minimal) sets of components which are deemed to be faulty and that can be substituted for *repair*, in the second type of diagnosis the underlying goal is to diagnose the 'disease' in order to provide the right *remedy* (that often cannot be a substitution). For these reasons, considering the social system domain, it makes sense to deal not only with positive, normal institutional models (e.g. buyer and seller in a sale contract), but also with explicitly faulty ones (e.g. tax evaders).

Despite these differences, however, abductive diagnosis and consistency-based diagnosis have been recognized as two poles of a spectrum of types of diagnosis [10]. In effect, we find contributions extending consistency-based diagnosis with faulty models (e.g. [15]) and abductive diagnosis with models of correct behaviour. In a more principled way, [25] shows that the two types of diagnosis can be unified relying on a *stable model semantics* (the same used in ASP), essentially because it considers the distinction and separate treatment of *strong negation* and *negation as failure*.

2.0.4 Deciding additional investigations

During a diagnostic process, it is normal to consider the possibility of conducting additional investigations (measurements, in the case of electronic devices) in order to conclusively isolate the set of faulty components, or more generally, to reduce the set of hypothetical explanations. For simplicity, we will neglect this aspect in this work; for completeness, however, we highlight two main directions investigated in the literature. The most frequently used approach, first proposed in [15], is to use a *minimum entropy* method to select which measurement to do next: choosing the datum which minimizes the entropy of the candidate after the measurement is equivalent to deciding the source that provides the maximum *information* to the diagnoser (cf. [**?**]). As this method considers only one additional source per step, it is also called *myopic*. The second approach proposes instead *non-myopic* or *lookahead* methods, i.e. deciding multiple steps to be performed at once, see e.g. [**?**]. In principle, this is the way to proceed when we account strategies for collecting information to minimize or control side-effects.

2.1 Diagnosis of Multi-Agent Systems

The association of diagnosis with *multi-agent systems* (MAS) is not very common in the literature, although the number of studies is increasing. In general, contributions alternatively refer to only one of the two natures of MAS, i.e. mechanism of distributed computation or framework for the instantiation of agent-based models. Therefore, on one side, MAS are proposed as a solution to perform diagnosis of (generally non-agent) systems, like in [27, 24]. On the other side, understanding of social failures is expressed as a problem of social coordination—see for instance [20, 19]. Unfortunately, the latter have generally a design-oriented approach, consequently, non-compliance and social failures are seen has a design issue, rather than systemic phenomena, as would be in a "natural" social system. For this reason, they share a perspective similar to works on checking non-compliance at regulatory level, e.g. [16, 18]: system (normative) requirements are literally taken as the reference on which to test compliance of business processes. Unfortunately, in doing this, we are not able to scope behaviours that superficially look compliant, but, for who knows the 'game', they are not.

Using agent-roles instead of roles The idea of using normative sources is related to the *role* construct; agents are usually seen as enacting certain institutional/organizational roles (e.g. [12]), inheriting their normative characterization. An alternative approach, from which this contribution stems out, has been proposed in [3], constructed on *agent-role* models: constructs that include the coordination of roles. The agent-role model share elements with those used in *intention-recognition* studies, and in particular with those based on logic approaches—see [28] for an overview—grown out from traditional AI accounts of story understanding and abduction. However, from a conceptual point of view, the "first principles" we are considering with agent-roles are not simple rules, but knowledge structures building upon practical reasoning constructs [34] and institutional positions [33]. More importantly, agent-roles are defined not only by a *script*, but also by a *topology*. By allowing to have multiple identities distributed on the topology, the agent-role model enable to take into account the existence of *collective agencies*, transcending the individual social participants.

3 CASE STUDY: SWAP SCHEMES IN REAL-ESTATE TRANSACTIONS

In the following section, we will focus on a well-known type of real-estate fraud, of the family of *swap schemes*, and present a few similar prototypical patterns. In a market context, a swap scheme establishes coordinations between dual groupings of buyers and sellers; as these parties are expected to compete within that institutional framework, it essentially undermines the *arm's length* principle of the market. On small economic scale this is not forbidden: e.g. "if you make me pay less for the guitar that your father is selling, I would make you pay less for my brother's motorcycle." However, in real-estate transactions, property transfer taxes apply. The full interaction includes the tax administration, and in these conditions swap schemes become means to reduce the amount of taxes due and, therefore, are not permitted.

3.1 Outline of a database of scenarios

Let us consider a simplified real estate market, with economic actors buying and selling houses of type A and of type B. Property transfer tax is 6% of the sale price, and the buyer and the seller have both nominally the burden to pay it (the actual distribution amongst the parties is however not fixed a priori). Besides the normal sale, we take into account three different scenarios: a swap scheme implementing a real-estate fraud, a hidden payment, and a wrong appraisal.

Example 1 (REAL ESTATE FRAUD, SWAP SCHEME)**.** *X and Y wants to exchange their properties: X owns a real estate of type A; Y owns one of type B, both worth €10 million. Instead of paying €600,000 per each in taxes, they set up reciprocal sales with a nominal price of €5 million, thus dividing the taxes due in half.*

The scheme is illustrated in Fig. 1. The picture highlights two coordination levels:

- an *intentional coordination* level, generally referring to some composition of institutional roles (in our case buyer/seller structures, the dashed boxes in the figure);
- a *scenario coordination* level, responsible of the synchronization of operations between the intentional coordination structures.

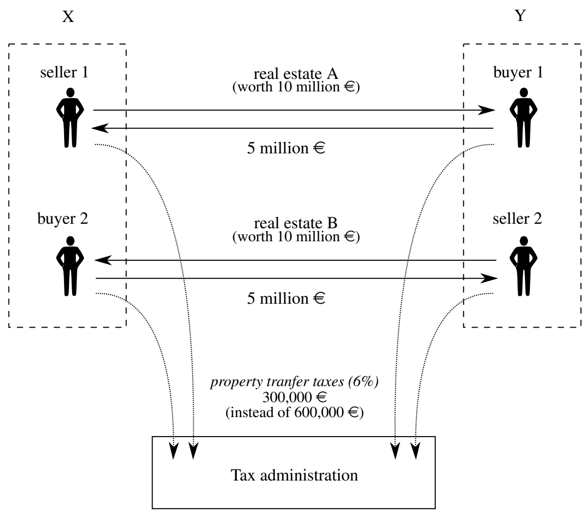

Figure 1. Topology of a real estate fraud based on a swap scheme

The first is the domain of *internal topologies* of agent-roles. The second is the domain of *coupling* configurations of agent-roles, i.e. of *external topologies*, specified as MAS.

The structures enabling coordination (at both levels) may be physical bodies, but also social bodies as natural, informal groupings of people (e.g. father and son), organizations (e.g. employer and employee), etc. It may be anything that suggests a sharing, a *concentration of interests*, or an existence of *stable inter-dependencies*, that may undermine the arm's length principle. At the scenario level, however, the relation is not necessarily as structured as the examples just given. In the case of bribery, for instance, there is typically no other relation between the parties beside a *contingent agreement*. Similarly, a swap-scheme may be performed by two real-estate agencies on a contingent basis.

Example 2 (HIDDEN PAYMENT). *X wants to give €300,000 to Y, and, as Y is also interested in X's house, X sells Y that house, worth €500,000, for €200,000.*

A hidden payment is usually economically advantageous for both parties because property transfer generally has lower taxation than other forms of transfer.

Example 3 (WRONG APPRAISAL). *X needs to sell his house. Not knowing the current prices for the area, he sells the house for €200,000 to Y, while at market price, the building would be worth around €500,000.*

4 OPERATIONALIZATION OF MONITORING AND DIAGNOSIS

In this exercise, we imagine taking the role of the tax administration, with the intent of monitoring the payment of taxes, possibly diagnosing (and also explaining) supposed institutional failures.[6] Note that the tax administration has only a *partial view* of the communications of the parties: in our simplified world, only sale declarations and tax payment receipts.

Types of failures The starting point of the operationalization is to collect the agent-roles of the domain relevant to the tax administration. The first set is given by simple intentional characterizations of *normal institutional roles*, i.e. buyers and sellers paying their taxes. From this, we can construct possible failure modes as violations of role obligations, dealing with representations of *negative events* (negative as they are defined by the failure of expectations concerning events). In this specific example, tax payment may be:

(i) completely missing, as failure to pay *tout court*,
(ii) wrong, as failure to pay the fixed amount of taxes (e.g. 6% of the sale price)
(iii) wrong, as failure to pay the 'right' amount of taxes, in terms of *reasonableness*, i.e. of what could have been expected to be paid to the tax administration for the sale of that property.

The third situation covers the case of swap-schemes or other tax evasion manœuvers; it is evidently more difficult to scope, as it requires an evaluation in terms of the social domain semantics—in this case, of the market pricing rationality. This is the domain in which the agent-role concept makes particularly the difference.

4.1 Monitoring

As we know that certain social participants may be non-compliant, we need to set up an adequate monitoring procedure. A first requirement of adequacy is the possibility of *discriminating* cases of non-compliance from those of compliance. This actually supports a general principle for choosing monitoring targets:

Proposition 1. *Outputs of* contrast *operations between compliant and non-compliant scenarios lead to identifying* events *or* threshold conditions *associated to suspicious transactions.*

The set of discriminating elements is constructed in terms of what is available through the monitoring, i.e. the 'perceptual' system of the agency. If the diagnostic agent is not able to monitor any discriminatory element, then the contrasting principle will not be exploitable and there will be no mean to recognize non-compliance. In our example, as the tax administration has direct access only to sale declarations and tax payment receipts, it is amongst these sources that we have to scope signs of potential failures.

Note that the contrast operation can be implemented thanks to the availability of executable models: by *executing* normal and failure models, we can predict the different traces they would produce, and then contrast them. In principle, however, we could refer directly to the traces. For instance, in medicine, failure modes are usually directly associated to symptoms, without explaining why a certain disease produces these symptoms. In the general case, however, this solution has limitations, as it assumes a relative invariance of the chain of *transmission* going from the source phenomenon to the perceptual system of the observer, which is not granted in a social system. Considering explicitly the underlying behavioural mechanism allows us to deal separately with such 'transmission' component.

We apply the previous principle to the three types of negative events. Case (i) requires the implementation of a *timeout* mechanism that asynchronously triggers the failure. Case (ii) requires a check *synchronously* to the receipt of payment; it can be implemented with a simple operational rule. Case (iii) is more complex: to conclude

[6] It is worth to observe that compliance and non-compliance are qualifications relative to the position of the diagnostic agent in the social system. For instance, in a world of liars, truth-tellers would fail in respect to the social practice of systematically lying.

that a price is reasonable requires us to assess the market price of that property, and to decide what deviation from market price is still acceptable. Let us arbitrarily specify this deviation as 40% of the market price, knowing that statistical methods may suggest more appropriate values. Therefore, the price provided in the sale declaration can be taken as a threshold to consider a certain sale price as *suspicious*. If implemented in Prolog, the qualification rule would look like the following code:

```
suspiciousPrice(Price, Estate, Time) :-
  marketPrice(MarketPrice, Estate, Time),
  Price =< (MarketPrice * 60)/100.

suspiciousSale(Seller, Buyer, Estate, Price, Time) :-
  declaration(sale(Seller, Buyer, Estate, Price, Time)),
  suspiciousPrice(Price, Estate, Time).
```

Clearly, this is a simple case. In general, multiple factors may concur with different weight to increase the suspiciousness of transaction.

In absence of average market price As we confirmed from talking with experts of the tax administration, the practical discrimination used by investigators to discover potential tax frauds is actually built upon comparisons with average market prices. Unfortunately, average market prices are not easy to be access in reality and, when they are, they may be not representative for that specific case.[7] A first solution would then be to refer to domain experts, e.g. appraisal agents, but these externalizations, where available, obviously increase the costs of investigation. A simple way to overcome the problem of assessing the market price of a certain real-estate property is to check the value of the same real-estate in previous sale transactions. In the case of swap schemes, the new owners tend to sell the recently acquired property after a relatively short time, but for a much higher price, even in the presence of relatively stable prices. From an operational point of view, this would correspond simply to a different tracking of the suspiciousness relation.

4.1.1 Diagnosis

When identified, suspicious transactions should trigger a diagnostic process in order to establish *why* the failure occurred. In general, the same 'symptoms' may be associated to diagnostic and non-diagnostic explanations. For instance, going through the known scenarios, a low price in a sale transaction may be due not only to a swap scheme, but also to a hidden payment, or it may simply be due to an error in the appraisal of the estate by the offeror. Interestingly, even if plausible, wrong appraisal is not taken into account by the tax administration. Why? Evidently, this choice is determined by the *strict liability* of these matters[8], but it may be seen as a consequence of a more fundamental issue: the tax administration cannot possibly read the mind of offeror to check the veracity of his declaration. A price that is not 'reasonable' cannot but be interpreted as an *escamotage* of both parties to avoid or reduce the tax burden.

Direct diagnostic mechanism In a simplistic form, direct evidence for a supposed swap-scheme would consist of two sets of buyers and sellers that have performed suspicious sales:

```
actionEvidenceOfSwap(
  sale(Seller1, Buyer1, EstateA, PriceA, Time1),
  sale(Seller2, Buyer2, EstateB, PriceB, Time2)
  ) :-
  suspiciousSale(Seller1, Buyer1, EstateA, PriceA, Time1),
  suspiciousSale(Seller2, Buyer2, EstateB, PriceB, Time2),
  not(EstateA = EstateB),
  not(Seller1 = Seller2), not(Buyer1 = Buyer2).
```

This is however not sufficient: sellers and buyers may have performed these transactions independently, and therefore this evaluation doesn't consider minimal *circumstantial* elements to support a swap-scheme rather than e.g. two hidden payments. In order to overcome this problem, we have to take into account explicitly a *relatedness* condition.

```
actionAndCircumstantialEvidenceOfSwap(
  sale(Seller1, Buyer1, EstateA, PriceA, Time1),
  sale(Seller2, Buyer2, EstateB, PriceB, Time2)
  ) :-
  actionEvidenceOfSwap(
    sale(Seller1, Buyer1, EstateA, PriceA, Time1),
    sale(Seller2, Buyer2, EstateB, PriceB, Time2)
  ),
  relatedTo(Seller1, SharedStructure1),
  relatedTo(Buyer2, SharedStructure1),
  relatedTo(Seller2, SharedStructure2),
  relatedTo(Buyer1, SharedStructure2).
```

An example of *relatedness* condition between buyer and seller may be, for instance, their participation in a common social structure (family, company, etc.), that may place its members outside the arm's length principle of the market. This condition acknowledges *potential* intentional coordination, i.e. a plausible concentration of *interests* that makes the transaction definitively suspect.[9]

The existence of a coordination structure at the scenario level, i.e. between such shared structures, would be additional evidence, but it is not necessary, as the scheme may be performed on a contingent basis (§ 3.1). Interestingly, the 'hidden payment' case turns out to be a minimal version of a swap-scheme:

```
actionAndCircumstantialEvidenceOfHiddenPayment(
  sale(Seller, Buyer, Estate, Price, Time)
  ) :-
  suspiciousSale(Seller, Buyer, Estate, Price, Time),
  relatedTo(Seller, SharedStructure),
  relatedTo(Buyer, SharedStructure).
```

By extension, we could imagine swap-schemes implemented through *networks* of buyer and sellers. This would be, for instance, a simple diagnostic test for swap-schemes performed on three-node networks:

```
actionAndCircumstantialEvidenceOf3Swap(
  sale(Seller1, Buyer1, EstateA, PriceA, Time1),
  sale(Seller2, Buyer2, EstateB, PriceB, Time2))
  sale(Seller3, Buyer3, EstateC, PriceC, Time3)
  ) :-
  suspiciousSale(Seller1, Buyer1, Estate1, PriceA, Time1),
  suspiciousSale(Seller2, Buyer2, Estate2, PriceB, Time2),
  suspiciousSale(Seller3, Buyer3, Estate3, PriceC, Time3),
  not(EstateA = EstateB),
  not(Seller1 = Seller2), not(Buyer1 = Buyer2),
  not(EstateB = EstateC),
  not(Seller2 = Seller3), not(Buyer2 = Buyer3),
  not(EstateA = EstateC),
  not(Seller1 = Seller3), not(Buyer1 = Buyer3),
  relatedTo(Seller1, SharedStructure1),
  relatedTo(Buyer3, SharedStructure1),
  relatedTo(Seller2, SharedStructure2),
  relatedTo(Buyer1, SharedStructure2),
  relatedTo(Seller3, SharedStructure3),
  relatedTo(Buyer2, SharedStructure3).
```

The inclusion of a third element breaks the direct connection between the initial parties, but the code makes explicit the pattern that can be extended by induction. More formally:

[7] On the one hand, prices of real estate properties in public offers often do not correspond to the actual prices of sale. On the other hand, the heterogeneity of real estate properties, the imperfect alignment between cadastral information and real situations, the dynamics of value associated to neighbourhoods and other relevant factors make it difficult to consider as reliable the application of average measures on actual cases.

[8] See note 3.

[9] This is evidently similar to the issue of *conflict of interest*: a person in power may be in a situation in which his discretion to reach the primary intents defined by his role may be biased towards the achievement of other intents.

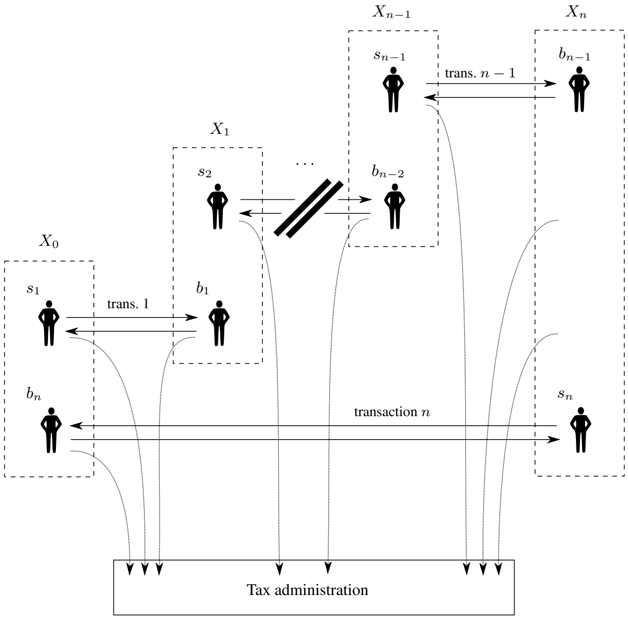

Figure 2. Swap scheme with n nodes.

Definition 1 (GENERALIZED SWAP-SCHEME THROUGH SALES)**.** Given n sale transactions, naming b_i and s_i respectively the buyer and the seller of a transaction i, a swap scheme holds if the following relatedness relations are established:

- between s_1 and b_n (named X_0)
- with $0 < i \leq n$, between s_i and b_{i-1} (named X_i)

The associated topology is illustrated in Fig. 2. It would certainly be interesting to evaluate mechanisms like this on data sets such as those released with the so-called Panama papers.

4.2 Improving the reasoning mechanism

The diagnostic mechanism proposed here leverages the advantages of *backward chaining* given by Prolog, i.e. of reasoning opportunistically in order to reach a conclusion about a certain epistemic goal. In a way, this is an opposite solution than the operationalization we proposed in *explanation-based argumentation* (EBA) [31], based on ASP, where factors brought by the observation are used to allocate *all* possible scenarios. On the other hand, it suffers from two important limitations. First, it relies on a *closed-world assumption* (CWA), i.e. *negation as failure* is automatically interpreted as *strong negation*. Second, it requires an explicit query to trigger the inferential process, but, in a practical setting, the monitoring and diagnostic process should be reactive to the reception of new observations. Therefore, a more plausible monitoring mechanism should look like the following *event-condition-action* (ECA) rule:

(E) when you receive a declaration,
(C) if it is suspicious,
(A) trigger the diagnostic process.

Third, the diagnostic process should consider the whole family of scenarios that are associated to that 'symptom', and should consider that there may be *missing information*. One way to proceed in this respect is to integrate a solution similar to EBA, i.e. of generating at need potential scenarios. Relevant known facts are used to fill fit scenarios belonging to this family, pruning impossible (according to logic constraints), or implausible (according to prior commitments) ones. Note that this family can be compiled *offline*, as much as the discriminatory power of the different factors allow. This information may be used to lead the investigation steps to be acted upon in real-time.

In this scenario, the procedural aspect was not essential, but in general, it may be.In related works, for instance, we built our models using (extensions of) Petri net [30, 32]. Petri net can be mapped to logic programming using for instance Event Calculus [29] or similar techniques; this can be related to *composite event recognition* approaches (e.g. [1]) suggest the use of intermediate caching techniques to improve the search. Another solution would be to instead maintain the process notation, and compute fitness decomposing the family of scenario in a hierarchy of *single-entry-single-exit* (SESE) components (e.g. [23]).

4.2.1 *Computational complexity*

Model-based diagnosis (MBD) is known to be a hard computational problem, namely exponential to the number of components of the diagnosed systems (see e.g. [4]). For this reason, diagnostic algorithms traditionally focus on minimal diagnoses, i.e. of minimal cardinality (involving minimal subset of faulty components), an approach that is also known as the *principle of parsimony* [26]. This principle is not directly applicable to our framework, as the system components are not agent-players, but agent-roles enacted by agent-players; each component is therefore 'invisible' to the observation, and can be tracked only as a mechanism involving individual elements.

Fortunately, it has been shown that the exponential increase of computational burden may still be reduced using a mixture of decomposition techniques and statistical information. In this chapter, we have overlooked this problem, as we focused on justifying the proposed method providing a working example of an application. We can, however, trace next directions to investigate. As we said in the previous section, the family of scenarios associated to a certain alarming event is known in advance. Therefore, some knowledge compilation techniques may produce important advantages, deriving heuristic knowledge for heuristic problem-solvers, without restarting from first principles (e.g. [5, 9]). Statistical information may instead be used to focus only on a limited set of most probable *leading* hypothesis [15]. It has been also suggested to control complexity by using hierarchical models, i.e. models with different levels of abstraction [22, 6, 35]. This is in principle directly possible with agent-roles. All these aspects remain to be investigated.

5 CONCLUSION

As already stated in the title, this paper is meant to describe an exercise of computational implementation, targeting a specific problem, exploiting part of the conceptual framework presented in previous works [2, 3]. For reasons of opportunity, we neglected many other practical and theoretical aspects that have been investigated in parallel, and that should be taken into account to get the full picture. For instance, about the *representation* of agent-roles, we have identified in *positions* the fundamental components, defined respectively towards another party for normative functions, in the tradition of Hohfeld's analytic framework [33], and towards the environment for practical reasoning purposes [34]. We have investigated the *acquisition* of agent-roles starting from UML-like diagrams [30] and from interpretations of narratives [32]. In these works we worked with (extensions of) Petri nets, also in order to set a natural convergence to the usual notation used for business process models.

On the other hand, this simplification allowed to appreciate instead the problems of settling a real-time model-based diagnosis activity in operations. It is easy to imagine further developments from the insights gained from this exercise. We will just name a few of them: a formalization of the *contrast* operation; the 'compilation' of the collected scenarios in knowledge bases optimized for monitoring and for diagnosis; the interface of EBA with backward-chaining, in order to take into account competing scenarios and the possibility of missing information; the possibility of composing multiple scenarios via planning, taking into account diversional behaviours (this would not be possible with diagnostic systems not relying on models); an investigation on the resulting computational complexity.

REFERENCES

[1] Alexander Artikis, Marek Sergot, and Georgios Paliouras, 'An Event Calculus for Event Recognition', *IEEE Transactions on Knowledge and Data Engineering*, **27**(4), 895–908, (2015).

[2] Alexander Boer and Tom van Engers, 'An agent-based legal knowledge acquisition methodology for agile public administration', in *Proceedings of the 13th International Conference on Artificial Intelligence and Law - ICAIL '11*, pp. 171–180, New York, (2011). ACM Press.

[3] Alexander Boer and Tom van Engers, 'Diagnosis of Multi-Agent Systems and Its Application to Public Administration', in *Business Information Systems Workshops*, volume 97 of *Lecture Notes in Business Information Processing*, pp. 258–269. Springer, (2011).

[4] Tom Bylander, Dean Allemang, Michael C. Tanner, and John R. Josephson, 'The computational complexity of abduction', *Artificial Intelligence*, **49**, 25–60, (1991).

[5] B. Chandrasekaran and Sanjay Mittal. Deep versus compiled knowledge approaches to diagnostic problem-solving, 1983.

[6] Luca Chittaro and Roberto Ranon, 'Hierarchical model-based diagnosis based on structural abstraction', *Artificial Intelligence*, **155**(1-2), 147–182, (may 2004).

[7] Luca Console and Oskar Dressier, 'Model-based diagnosis in the real world: Lessons learned and challenges remaining', *IJCAI International Joint Conference on Artificial Intelligence*, **2**, 1393–1400, (1999).

[8] Luca Console, Daniele Theseider Dupré, and Pietro Torasso, 'A theory of diagnosis for incomplete causal models', *Proceedings 11th International Joint Conference on Artificial Intelligence*, 1311–1317, (1989).

[9] Luca Console, Luigi Portinale, and Daniele Theseider Dupré, 'Using compiled knowledge to guide and focus abductive diagnosis', *IEEE Transactions on Knowledge and Data Engineering*, **8**(5), 690–706, (1996).

[10] Luca Console and Pietro Torasso, 'A spectrum of logical definitions of model-based diagnosis', *Computational Intelligence*, **7**(3), 133–141, (1991).

[11] P. T. Cox and T. Pietrzykowski, 'Causes for Events: Their Computation and Applications', in *Deductive Databases, Planning, Synthesis - 8th International Conference on Automated Deduction*, volume LNCS 230, pp. 608–621, (1986).

[12] Mehdi Dastani, M. Birna van Riemsdijk, Joris Hulstijn, Frank Dignum, and John-Jules Ch. Meyer, 'Enacting and deacting roles in agent programming', *AOSE 2004: Proc. of 5th Int. Workshop on Agent-Oriented Software Engineering*, (2004).

[13] Randall Davis. Diagnostic reasoning based on structure and behavior, 1984.

[14] Johan de Kleer and BC Brian C. Williams, 'Diagnosing multiple faults', *Artificial intelligence*, **32**(1987), 97–130, (1987).

[15] Johan de Kleer and Brian C. Williams, 'Diagnosis with behavioral modes', *International Joint Conference On Artificial Intelligence*, 1324–1330, (1989).

[16] Guido Governatori, 'Business Process Compliance: An Abstract Normative Framework', *it Information Technology*, **55**(6), 231–238, (2013).

[17] Edwin Hutchins, 'Enaction, Imagination, and Insight', in *Enaction: towards a new paradigm in cognitive science*, 425–450, MIT Press, Cambridge, Massachusetts, (2010).

[18] J I E Jiang, Huib Aldewereld, Virginia Dignum, and Yao-hua Tan, 'Compliance Checking of Organizational Interactions', *ACM Transactions on Management Information Systems*, **5**(4), 1–24, (2014).

[19] Özgür Kafal and Paolo Torroni, 'Exception diagnosis in multiagent contract executions', *Annals of Mathematics and Artificial Intelligence*, **64**(1), 73–107, (mar 2012).

[20] Meir Kalech, 'Diagnosis of coordination failures: a matrix-based approach', *Autonomous Agents and Multi-Agent Systems*, **24**, 69–103, (jul 2012).

[21] Peter J.F. Lucas, 'Analysis of notions of diagnosis', *Artificial Intelligence*, **105**(1-2), 295–343, (1998).

[22] Igor Mozetič, 'Hierachical Model-based Diagnosis', *International Journal of Man-Machine Studies*, **35**(3), 329—-362, (1991).

[23] Jorge Munoz-Gama, Josep Carmona, and Wil M P Van Der Aalst, 'Single-Entry Single-Exit decomposed conformance checking', *Information Systems*, **46**, 102–122, (2014).

[24] M. Pipattanasomporn, H. Feroze, and S. Rahman, 'Multi-agent systems in a distributed smart grid: Design and implementation', *2009 IEEE/PES Power Systems Conference and Exposition*, 1–8, (mar 2009).

[25] Chris Preist, Kave Eshghi, and Bruno Bertolino, 'Consistency-based and abductive diagnoses as generalised stable models', *Annals of Mathematics and Artificial Intelligence*, **11**(1-4), 51–74, (1994).

[26] Raymond Reiter, 'A theory of diagnosis from first principles', *Artificial Intelligence*, **32**(1), 57–95, (apr 1987).

[27] Nico Roos, Annette ten Teije, and Cees Witteveen, 'A protocol for multi-agent diagnosis with spatially distributed knowledge', *Proceedings of the second international joint conference on Autonomous agents and multiagent systems - AAMAS '03*, 655, (2003).

[28] Fariba Sadri, 'Logic-based approaches to Intention Recognition', in *Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives*, eds., N.-Y. Chong and F. Mastrogiovanni, 346–375, IGI Global, (2012).

[29] Murray Shanahan, 'The event calculus explained', *Artificial intelligence today*, 409–430, (1999).

[30] Giovanni Sileno, Alexander Boer, and Tom van Engers, 'From Inter-Agent to Intra-Agent Representations: Mapping Social Scenarios to Agent-Role Descriptions', in *Proc. 6th Int. Conf. Agents and Artificial Intelligence (ICAART 2014)*, (2014).

[31] Giovanni Sileno, Alexander Boer, and Tom van Engers, 'Implementing Explanation-Based Argumentation using Answer Set Programming', in *11th International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2014)*, (2014).

[32] Giovanni Sileno, Alexander Boer, and Tom Van Engers, 'Legal Knowledge Conveyed by Narratives: Towards a Representational Model', *Proceedings of the Workshop on Computational Models of Narrative (CMN 2014)*, 182–191, (2014).

[33] Giovanni Sileno, Alexander Boer, and Tom van Engers, 'On the Interactional Meaning of Fundamental Legal Concepts', *JURIX 2014: 27th Int. Conf. Legal Knowledge and Information Systems*, **FAIA 271**, 39–48, (2014).

[34] Giovanni Sileno, Alexander Boer, and Tom van Engers, 'Commitments, Expectations, Affordances and Susceptibilities: Towards Positional Agent Programming', *PRIMA 2015: 18th Conf. on Principles and Practice of Multi-Agent Systems*, **LNCS 9387**, 687–696, (2015).

[35] Roni Stern, Meir Kalech, and Orel Elimelech, 'Hierarchical Diagnosis in Strong Fault Models', *DX Workshop*, (2014).