Improving Model-Based Diagnosis through Algebraic Analysis: the Petri Net Challenge

Luigi Portinale
Dipartimento di Informatica - Universita' di Torino
C.so Svizzera 185 - 10149 Torino (Italy)
e-mail: portinal@di.unito.it

Abstract

The present paper describes the empirical evaluation of a linear algebra approach to model-based diagnosis, in case the behavioral model of the device under examination is described through a Petri net model. In particular, we show that algebraic analysis based on P-invariants of the net model, can significantly improve the performance of a model-based diagnostic system, while keeping the integrity of a general framework defined from a formal logical theory. A system called INVADS is described and experimental results, performed on a car fault domain and involving the comparison of different implementations of P-invariant based diagnosis, are then discussed.

Introduction

In some recent papers (Portinale 1993), we have shown that Petri nets (PNs) (Murata 1989) can be fruitfully employed to face the problem of model-based diagnosis. This is accomplished by taking into account a formal logical framework of reference, defining classical notions (from the AI point of view) concerning the characterization of a diagnostic problem. In particular, it is shown that classical reachability analysis of PNs can naturally be exploited in order to realize "formally correct" (with respect to the logical framework of reference) diagnostic inference procedures. In the present paper, we focus on the empirical evaluation of a particular reachability analysis technique, namely P-invariant analysis, in order to show its practical usefulness and its possible advantages with respect to a logical inference mechanism. This analysis exploits a matrix representation of the net model and it is grounded on a linear algebra algorithm able to compute the so-called P-invariants of the net. They informally represent the correspondent of logical derivations and form the basis for the computation of the diagnoses. We will report on the empirical results obtained from some tests performed on a car fault domain, by comparing different implementations of P-invariant diagnosis and a classical abductive approach.

Petri Nets: Outline

A Petri net is a directed bipartite graph $N = \langle P, T, F \rangle$ whose vertices are called *places* (the elements of P represented as small circles) and *transitions* (the element of T represented as bars). The set of arcs is represented by F. In case the transitive closure F^+ of the arcs is irreflexive, the net is said to be *acyclic*. In a Petri net, an arc multiplicity function is usually defined as $W : (P \times T) \cup (T \times P) \rightarrow I\!N$; in case W is such that $W(f) = 1$ if $f \in F$ and $W(f) = 0$ if $f \notin F$, the net is said to be an *ordinary Petri net*. We will mainly be interested in such a kind of nets. For each $x \in P \cup T$ we use the classical notations ${}^\bullet x = \{y / yFx\}$ and $x^\bullet = \{y / xFy\}$. If ${}^\bullet x = \emptyset$, x is said to be a *source*, while if $x^\bullet = \emptyset$, x is said to be a *sink*. A *marking* is a function from the set of places to nonnegative integers, represented by means of *tokens* into places; a place containing a token is said to be *marked*. A *marked Petri net* is a pair (N, μ) where N is a Petri net and μ is a marking. The dynamics of the net is described by moving tokens from places to places according to the *concession* and *firing rules*. In ordinary Petri nets we say that a transition t *has concession* at a marking μ if and only if $\forall p \in^\bullet t \ \mu(p) \geq 1$. If a transition t has concession in a marking μ, it may *fire* (execute) producing a new marking μ' such that $\forall p \in P \ \ \mu'(p) = \mu(p) - W(p,t) + W(t,p)$. A marking μ' is *reachable* from a marking μ in a net N ($\mu' \in R(N, \mu)$) if and only if there exists a sequence of transitions producing μ' from μ in N. If a place of a marked net cannot be marked with more than one token, the place is said to be *safe*; if the property holds for every place, the net itself and every marking are said to be *safe*. Given a Petri net $N = \langle P, T, F \rangle$, if $n = |T|$ and $m = |P|$, the *incidence matrix* of N is the $n \times m$ matrix of integers $A = [a_{ij}]$ such that $a_{ij} = W(i,j) - W(j,i) (i \in T, j \in P)$ An m-vector of integers Y such that $A \cdot Y = 0$ is said to be a *P-invariant* of the net represented by A, the entry $Y(j)$ corresponding to place j. The *support* σ_Y of a P-invariant Y is the subset of places correspond-

ing to nonzero entries of Y. In a dual way, if A^T is the transpose matrix of A, an n-vector of integers X such that $A^T \cdot X = 0$ is said to be a *T-invariant* (entries corresponding to transitions). It is well known that any invariant can be obtained as a linear combination of invariants having minimal (with respect to set inclusion) supports (Murata 1989).

Petri nets and Model-based Diagnosis

Model-based diagnosis deals with the problem of determining the explanation of the abnormal behavior of a given device, by reasoning on a model (usually a behavioral model) of such a device (Hamscher, Console, & de Kleer 1992). Approaches based on "consistency" between the observed and the predicted system behavior (with some components assumed to be faulty) are usually considered when the model represents the expected behavior of the system; however, when also the faulty behavior is taken into account, approaches based on "abduction" can be more adequately adopted.

Since both purely consistency based and purely abductive approaches suffer from some drawbacks, some effort has been done in order to combine them (Poole 1989; Console & Torasso 1991). In the present paper, we will refer to the framework defined in (Console & Torasso 1991).

Definition 1 *A model-based diagnostic problem is a tuple* $DP = \langle M, H, CXT, \langle \Psi^+, \Psi^- \rangle \rangle$ *where:*
- M is a logical theory representing the model of the system to be diagnosed;
*- H is a set of ground atoms of M identified as possible diagnostic hypotheses (*abducibles*);*
- CXT is a set of ground atoms of M representing contextual information;
- Ψ^+ is a set of ground atoms of M representing the observations to be covered in the current case;
- Ψ^- is a set of ground atoms of M representing the values of observable parameters conflicting with the observations.

We assume that M is represented by a set of definite clauses. This allows us to focus on a simple kind of model that is however representationally adequate for significant classes of behavioral models (see (Console *et al.* 1993)). If OBS is the set of current observations, the set Ψ^+ is in general a subset of OBS ($\Psi^+ \subseteq OBS$), while $\Psi^- = \{m(x)/m(y) \in OBS, x \neq y\}$. In a similar way, given the set CXT we define the set $CXT^- = \{c(x)/c(y) \in CXT\}$. The framework has the implicit assumption of abstracting from time; this allows us to further simplify the logical model by assuming M to be a set of definite clauses without recursion (i.e. a hierarchical definite logic program). We also assume the set OBS be composed by ground atoms having no consequences. Similarly, atoms in H cannot appear in the head of any clause (i.e. diagnostic hypotheses are independent) and so atoms in CXT.

Definition 2 *Given a diagnostic problem DP a diagnosis to DP is a set $E \subseteq H$ such that*
$\forall m(x) \in \Psi^+ \; M \cup CXT \cup E \vdash m(x)$
$\forall m(y) \in \Psi^- \; M \cup CXT \cup E \nvdash m(y)$

We will refer to a diagnostic problem defined in this way as a *logic-based diagnostic problem.* Notice that definition 2 does not require the set E to mention every abducible predicate of M; however, E could not be a *partial diagnosis* in the sense of (de Kleer, Mackworth, & Reiter 1992), since there could be an extension of E to unmentioned abducible predicates such that some atoms in the set Ψ^- are derived. However, in the following we will consider only fault models (i.e. models describing only the consequences of the faulty behavior of the device under examination); in this case, a diagnosis E is interpreted as assigning a "normal" or "correct" value to abducible predicates not mentioned in E. The capability of dealing with models mixing the correct and the faulty behavior of the system requires a slight revision of the definition of diagnosis, by considering the set Ψ^- to be a set of denials, to be used in a contrapositive way. This would allows us to get the equivalent of the *kernel diagnoses* defined in (de Kleer, Mackworth, & Reiter 1992)[1].

In (Portinale 1993) a simple Petri net model, called Behavioral Petri Net (BPN), has been introduced, in order to capture the representational issues discussed above.

Definition 3 *A Behavioral Petri Net (BPN) is a 4-tuple* $M = \langle P, T_N, T_{OR}, F \rangle$ *such that* $(P, T_N \cup T_{OR}, F)$ *is an acyclic ordinary Petri net that satisfies the following axioms:*
1. $\forall p \in P(|{}^\bullet p| \leq 1 \wedge |p^\bullet| \leq 1)$
2. $\forall p_1, p_2 \in P \; (({}^\bullet p_1 = {}^\bullet p_2) \wedge (p_1^\bullet = p_2^\bullet) \rightarrow p_1 = p_2)$
3. $\forall t \in T_N \; (|{}^\bullet t| = 1 \wedge |t^\bullet| > 0) \vee (|{}^\bullet t| > 0 \wedge |t^\bullet| = 1)$
4. $\forall t \in T_{OR}(|{}^\bullet t| \geq 2 \wedge |t^\bullet| = 1)$

The set of transitions is partitioned into two subset T_N and T_{OR}; those in the former set are the usual kind of transitions of ordinary Petri nets while a transition $t \in T_{OR}$ has concession in a marking iff at least one of its input places is marked. They are actually "macro-transitions" that can be obtained by means of a set of classical transitions (see (Portinale 1993)). It can be shown that a BPN models the same kind of knowledge of a hierarchical definite logic program. Figure 1 shows an example of a BPN corresponding to the following set of clauses (OR transitions are represented as empty thick bars):

$grcl(low) \wedge roco(poor) \rightarrow oils(holed) \qquad oils(holed) \rightarrow obca(huge\_am)$

[1]Notice that definition 2 can be directly used if we consider E to contain exactly one ground instance for each abducible predicate.

$roco(poor) \rightarrow jerk(very\_strong)$ $\quad osga(worn) \rightarrow oils(leaking)$
$oils(leaking) \rightarrow obca(small\_am)$ $\quad piws(worn) \rightarrow laoi(severe)$
$jerk(very\_strong) \rightarrow vibr(very\_strong)$ $\quad ente(incr) \rightarrow htin(red)$
$engi(run) \wedge laoi(severe) \rightarrow ente(incr)$ $\quad oils(holed) \rightarrow laoi(severe)$
$jerk(very\_strong) \wedge oils(leaking) \rightarrow laoi(severe)$

Table 1 shows the key for acronyms used. The net of figure 1 is just for explanatory purposes, corresponding to a simplified part of a more general model, describing the faulty behavior of a car engine. Notice that the BPN contains some "dummy places" (labeled in figure 1 with capital letters) used to split places representing ground atoms involved in the body of more than one clause. This allows us to identify the token flow on the net with logical derivations in the logical model. OR-transitions model alternative way of obtaining a given atom.

Since a BPN is acyclic, a partial order $\prec$ over transitions is defined as $t_1 \prec t_2 \leftrightarrow t_1 F^+ t_2$. A concession rule with priority for transitions can then be introduced, resulting in the *enabling rule* of a BPN.

Definition 4 *Given a BPN, a transition t is* enabled *(i.e. it may fire) in a given marking μ if and only if it has concession at μ and $\nexists t' \prec t$ such that t' has concession at μ.*

For example, in the net of figure 1, if both places *piws(worn)* and *oils(holed)* are marked, transition t_{27} is not enabled, since there is transition t_2 having concession and such that $t_2 \prec t_{27}$.

Definition 5 *An initial marking of a BPN is a safe marking μ_0 such that $\mu_0(p) = 1 \rightarrow^\bullet p = \emptyset$.*

A marked BPN is always considered with respect to a marking reachable from an initial marking. As shown in (Portinale 1993), every marked BPN is *safe* and in a marked BPN there is a unique marking, called the *final marking*, from which no transition can fire.

Given a BPN $N = \langle P, T_N, T_{OR}, F \rangle$ corresponding to a hierarchical definite logic program M, if B_M is

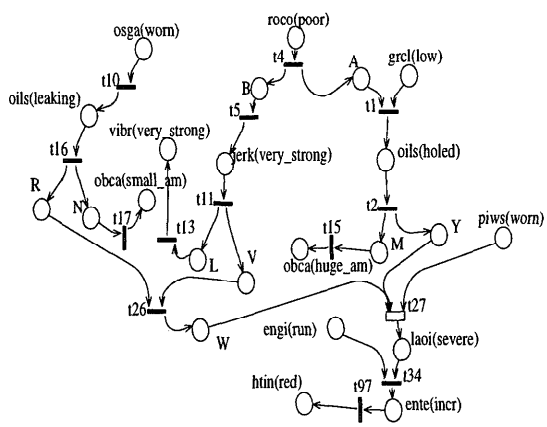

Figure 1: Example of a BPN

ENTITY	ACRONYM	ENTITY	ACRONYM
engine status	engi	engine temper.	ente
ground clear.	grcl	high temp. ind.	htin
jerks	jerk	lack of oil	laoi
oil below car	obca	oil sump gasket	osga
oil sump status	oils	piston wear	piws
road conditions	roco	vibrations	vibr

Table 1: Acronyms used in the BPN of fig. 1 and in the corresponding logical model

the Herbrand base of M, an *interpretation function* $\Phi : P \rightarrow B_M$ associating places of N with ground atoms of M can be defined. The function is in general a partial function; for example, in figure 1 the function Φ is considered undefined ($\perp$) for places labeled with capital letters (dummy places having no direct correspondence with ground atoms of M), while for the other places the label itself shows the value of Φ.

Given a conjunction of ground atoms J (represented as a set) we can determine a corresponding marking μ_J such that $\mu_J(p) = 1$ if $\Phi(p) \in J$ and $\mu_J(p) = 0$ otherwise.

Definition 6 *Given a logic-based diagnostic problem $DP = \langle M, H, CXT, \langle \Psi^+, \Psi^- \rangle \rangle$ and a BPN N_M corresponding to M, we can define the diagnostic problem in terms of the BPN model in the following way: BPN-DP= $\langle N, P_H, P_C, \langle P^+, P^- \rangle \rangle$ where $P_H = \{p \in P / \Phi(p) \in H\}$, $P_C = \{p \in P / \Phi(p) \in CXT\}$, $P^+ = \{p \in P / \Phi(p) \in \Psi^+\}$ and $P^- = \{p \in P / \Phi(p) \in \Psi^-\}$.*

Notice that, $\forall p \in P^+ \cup P^- \rightarrow p^\bullet = \emptyset$ (i.e. p is a sink place); similarly, $\forall p \in P_H \cup P_C \rightarrow^\bullet p = \emptyset$ (i.e. p is a source place). The formal connection between logic-based and BPN-based characterizations is established by means of the following theorem whose proof can be found in (Portinale 1993):

Theorem 1 *Given a logic-based diagnostic problem $DP = \langle M, H, CXT, \langle \Psi^+, \Psi^- \rangle \rangle$, let N_M be the BPN corresponding to M and μ_E^{CXT} be the marking corresponding to $E \cup CXT$ ($E \subseteq H$); if*
$(N_M, \mu_E^{CXT}) \vdash \alpha(c) \equiv$
$\exists \mu \in R(N_M, \mu_E^{CXT}) / \mu(p) = 1 \wedge \Phi(p) = \alpha(c)$
then $M \cup E \cup CXT \vdash \alpha(c) \leftrightarrow (N_M, \mu_E^{CXT}) \vdash \alpha(c)$

Definition 7 *Given a diagnostic problem BPN-DP= $\langle N_M, P_H, P_C, \langle P^+, P^- \rangle \rangle$, a* candidate diagnosis *is a marking μ_0 such that $\mu_0(p) = 1 \rightarrow p \in P_H$.*

We indicate with μ_C the marking corresponding to contextual information (i.e. $\mu_C(p) = 1 \leftrightarrow p \in P_C$) and with P_C^- the set of places corresponding to CXT^- (i.e. $P_C^- = \{p \in P / \Phi(p) \in CXT^-\}$).

Definition 8 *Given a diagnostic problem BPN-DP= $\langle N_M, P_H, P_C, \langle P^+, P^- \rangle \rangle$ a candidate diagnosis μ_0 is a solution to BPN-DP (i.e. is a diagnosis) if and only if*
$\forall p \in P^+ \ (N_M, \mu_0 \cup \mu_C) \vdash \Phi(p)$
$\forall q \in P^- \ (N_M, \mu_0 \cup \mu_C) \nvdash \Phi(q)$

Definition 9 *A marking μ of a Behavioral Petri Net* covers *a set of places Q if and only if $\forall p \in Q \rightarrow \mu(p) = 1$, while it* zero-covers *Q if and only if $\forall p \in Q \rightarrow \mu(p) = 0$.*

The following theorem provides us with an operational notion of diagnosis in a BPN framework (see (Portinale 1993) for the proof):

Theorem 2 *A candidate diagnosis μ_0 is a solution to BPN-DP= $\langle N_M, P_H, P_C, \langle P^+, P^- \rangle\rangle$ (i.e. is a diagnosis) if and only if the final marking μ of $(N_M, \mu_0 \cup \mu_C)$* covers *P^+ and* zero-covers *P^-.*

This means that the problem of finding the solutions to a diagnostic problem can be re-formulated as a reachability problem on the net model; this can classically be tackled in two different ways, with a *reachability graph approach* (as shown in (Anglano & Portinale 1994)) or with an *algebraic (invariant-based) approach.* The aim of this paper is to concentrate on invariant analysis and to discuss the performance of a diagnostic algorithm based on such a principle, with respect to a classical approach based on symbolic manipulation.

Diagnostic Reasoning by Computing P-Invariants

In this section, we will show how to generate an initial marking satisfying the condition of theorem 2 from a set of P-invariant supports. By definition, P-invariants of a net $N = \langle P, T, F\rangle$ correspond to T-invariants of its dual net $N_D = \langle T, P, F\rangle$. The following lemma has been proved in (Peterka & Murata 1989).

Lemma 1 *Let $N = \langle P, T, F\rangle$ be a Petri net such that $\forall t \in T |t^\bullet| \leq 1$ and $t \in T$ be a sink transition; there exists a T-invariant X of N such that $X(t) \neq 0$ if and only if t is firable from the empty marking.*

This means that in N there are some source transitions firing from the empty marking, eventually leading to the firing of t. Consider now an ordinary Petri net: in order a place p to be marked, there must be a transition $t \in^\bullet p$ that fire, while in order a transition t to fire, every place $p \in^\bullet t$ must be marked. If every transition of a Petri net has exactly one input place, the sentence "a place is marked" corresponds to the sentence "a transition can fire" in the dual net. Let us then consider the following transformation on a BPN:

$\wedge$-fusion. Given a BPN $N = \langle P, T_N, T_{OR}, F\rangle$, produce the ordinary Petri net $N' = \langle P', \{T_N \cup T_{OR}\}, F'\rangle$ as follows: for each $t \in T_N$ such that $^\bullet t = \{p_1, \ldots p_k\}$ $(k > 1)$ substitute in P the set $\{p_1, \ldots p_k\}$ with the place $p_{1,k}$ such that $^\bullet p_{1,k} = \bigcup_{i=1}^{k} {}^\bullet p_i$ and $p_{1,k}^\bullet = \{t\}$

This transformation simply collapses places that are "AND-ed" into a single place representing their conjunction; even if the resulting net is no longer a BPN, it encodes the same kind of knowledge of the original BPN. In fact, let us consider the interpretation function Φ of N and the following operator $\oplus$ on it:

$$\Phi(p) \oplus \Phi(q) = \Phi(p) \cup \Phi(q)$$

With $\Phi(p) \oplus \perp = \perp \oplus \Phi(p) = \Phi(p)$

We can define an interpretation function Φ' for N' from the interpretation function Φ of N as follows:

$$\Phi'(p) = \begin{cases} \Phi(p) & \text{if } p \in P \cap P' \\ \bigoplus_{i=1}^{k} \Phi(p_i) & \text{if } p = p_{1,k} \in P' - P \end{cases}$$

Figure 2 shows the net obtained from the BPN of figure 1 by means of the $\wedge$-fusion. Places $grcl(low)$ and A are collapsed into place "$grcl(low) + A$", V and R into place "$V + R$", $laoi(severe)$ and $engi(run)$ into place "$laoi(moder) + engi(run)$". The interpretation function Φ' is such that
$\Phi'(grcl(low) + A) = \{grcl(low)\}$, $\Phi'(V + R) = \perp$, $\Phi'(laoi(severe) + engi(run)) = \{laoi(severe), engi(run)\}$, $\Phi' \equiv \Phi$ for remaining places.

Theorem 3 *Given a BPN N_M corresponding to a hierarchical definite logic program M, let N'_M be the net obtained from N_M through the $\wedge$-fusion transformation, Φ' the interpretation function of N'_M and p a sink place; the following are equivalent propositions:*
1) there is a P-invariant Y of N'_M such that $Y(p) \neq 0$;
2) by marking source places p_s such that $Y(p_s) \neq 0$, the place p can eventually be marked;
3) $\bigcup_{p_s} \Phi'(p_s) \cup M \vdash \Phi'(p)$

Proof. 1) $\equiv$ 2) is a consequence of lemma 1 and of the fact that T-invariants of a net are P-invariants for

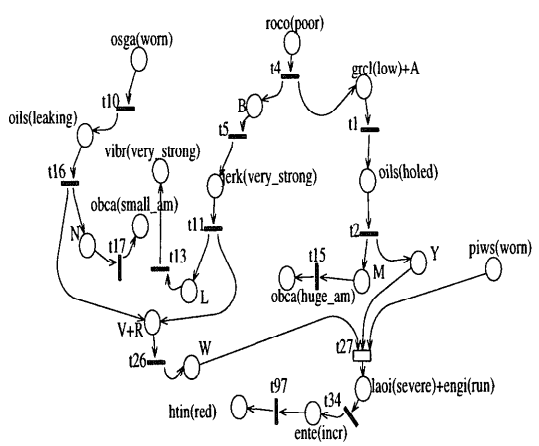

Figure 2: BPN for P-invariant Computation

its dual net. 2) $\equiv$ 3) is a consequence of theorem 1. □
From theorem 3 we conclude that the supports of the P-invariants of N'_M characterize the logical derivations from atoms representing diagnostic hypotheses and contexts, to atoms representing observable parameters. Consider for instance the following diagnostic problem BPN-DP= $\langle N_M, P_H, , P_C, \langle P^+, P^- \rangle \rangle$ where N_M is the net of figure 1 (we remember that such a net is intended to represent a fault model). Let us suppose to have the following set of observations:
$OBS = \{htin(red), obca(small\_am), vibr(normal)\}$
(i.e. the temperature indicator is red, there is a small amount of oil below the car and vibrations are normal). Contextual information are $CXT = \{grcl(normal), engi(run)\}$ (i.e. we are considering a car with a normal ground clearance and in the context of the engine being running). Let us also suppose that all the "abnormal" observations have to covered, then:
$P_H = \{piws(worn), osga(worn), roco(poor)\}$,
$P_C = \{engi(run)\}$ $\quad P_C^- = \{grcl(low)\}$
$P^+ = \{htin(red), obca(small\_am)\}$,
$P^- = \{vibr(very\_strong), obca(huge\_am)\}$.
The net of figure 2 is the result of the $\wedge$-fusion of N_M; the minimal supports of its P-invariants are:
$\sigma_1 = \{grcl(low) + A, oils(holed), roco(poor), M, obca(huge\_am)\}$
$\sigma_2 = \{grcl(low) + A, oils(holed), roco(poor), Y, ente(incr)$
$laoi(severe) + engi(run), htin(red)\}$
$\sigma_3 = \{B, roco(poor), jerk(very\_strong), osga(worn), V + R, W,$
$oils(leaking), laoi(severe) + engi(run), ente(incr), htin(red)\}$
$\sigma_4 = \{B, roco(poor), jerk(very\_strong), L, vibr(very\_strong)\}$
$\sigma_5 = \{obca(small\_am), osga(worn), oils(leaking), N\}$
$\sigma_6 = \{piws(worn), laoi(severe)+engi(run), ente(incr), htin(red)\}$
Consider for instance σ_4: we notice that the support contains the source place $roco(poor) \in P_H$ and the sink place $vibr(very\_strong) \in P^-$. From theorem 3 we conclude that $M \cup \{roco(poor)\} \vdash vibr(very\_strong)$; this means that any initial marking having place $roco(poor)$ marked is not a diagnosis, since it will eventually produce a final marking having place $vibr(very\_strong) \in P^-$ marked.

From these considerations, we can devise a P-invariant based diagnostic algorithm: after having computed the minimal supports of P-invariants, (efficient algorithms exist for this task (Martinez & Silva 1982)) those related to predictions corresponding to places in P^- are eliminated by taking into account the fact that if $\hat{\sigma}$ and $\hat{\sigma}'$ are two sets of ground atoms such that $\hat{\sigma} \subseteq \hat{\sigma}'$, if $\hat{\sigma} \vdash \alpha$ then $\hat{\sigma}' \vdash \alpha$; at the same way, supports containing places belonging to P_C^- are also eliminated. We have then to consider the coverability of P^+; for each place $p \in P^+$, we build from remaining supports the list of places having interpretation function corresponding to a diagnostic hypothesis and supporting p (i.e. contained in a P-invariant support containing p). Final diagnoses are obtained by combining such lists.

Let us consider again the diagnostic problem introduced above; supports σ_1, σ_2 are discarded since they contain place $grcl(low) + A$ such that $\Phi'(grcl(low) + A) = grcl(low) \in CXT^-$ (σ_1 also contains $obca(huge\_am) \in P^-$) and support σ_4 because it contains place $vibr(very\_strong) \in P^-$, Moreover, also support σ_3 is discarded because of the pruning of σ_4; indeed, $\hat{\sigma}_4 = \{roco(poor)\}$ and $\hat{\sigma}_3 = \{roco(poor), osga(worn), engi(run)\}$. Since $\hat{\sigma}_4 \subset \hat{\sigma}_3$, we need to prune also σ_3. Only supports σ_5 and σ_6 survive to the pruning phase and we then obtain:
$\hat{\sigma}_6 = \{osga(worn\}$ for $obca(small\_am) \in P^+$
$\hat{\sigma}_9 = \{piws(worn), engi(run)\}$ for $htin(red) \in P^+$
The only possible combination in this case is $\hat{\sigma}_6 \cup \hat{\sigma}_9 = \{piws(worn), engi(run), osga(worn)\}$ representing the diagnosis "$piws(worn) \wedge osga(worn)$" in the context "$engi(run) \wedge grcl(normal)$" (i.e. if the engine is running and the car has a normal ground clearance, the normal intensity of vibrations, the red temperature indicator and the small amount of oil below the car are explained by the fact that both the state of the pistons and the oil sump gasket are worn).

Experimental Results

We implemented the P-invariant approach to diagnosis in a system called INVADS (INVAriant based Diagnostic System) and we performed different series of experiments addressing the following two issues:
1. comparison of different implementations of invariant-based diagnosis;
2. comparison between invariant-based diagnosis and logical-abductive diagnosis.
Both types of experiments have been done on a BPN relative to a knowledge base describing the fault "causal" model of a car engine and consisting in more than 100 places and more than 100 transitions. We ran the experiments on a SUN Sparc station Classic with 32 Mbytes of memory; the software environment has been realized in SICStus prolog, with an embedded module for invariant computation written in C. We considered 48 different cases of car engine malfunctions in such a way to consider all the main fault evolutions described in the model. Different running of the same batch of cases have been considered for each implementation; results about the running time showed a quite low variance between different runs, so they have been simply averaged.

Implementation Testing

We tested three different kinds of implementation of P-invariant based diagnosis that we classified as follows:

off-line invariant computation (OFF); net simplification (SIM); observation addition (ADD).
The first kind of implementation (OFF) simply consists in the off-line computation of all the P-invariants of the net obtained from the $\wedge$-fusion on the BPN under examination; this approach makes explicit the information related to the P-invariants once for all and any diagnostic case that will be provided to the system, will use the same set of P-invariants. However, we have to search the solution in a search space (the set of P-invariants supports) that contains information that is not relevant to the current case. The complexity of a diagnostic algorithm based on this principle is just the complexity of the phase concerning the generation of diagnoses from P-invariant supports.

The SIM approach consists in simplifying the net with respect to the observations made in the case to be solved. This can be done by considering the sets P^+ and P^- from the current diagnostic problem, iteratively repeating the following actions, until no transition is removed.
for each place $p \notin P^+ \cup P^-$ **do remove** p;
for each transition $t/t^\bullet = \emptyset$ **do remove** t;
This allows us to consider only the part of the net relevant to the current set of observations and to compute the P-invariants only for this reduced net. A diagnostic algorithm based on SIM must take into account three different phases for each case to be solved: *net simplification*, *P-invariant computation* and *diagnosis generation*. Since the set of P-invariant supports from which to generate diagnoses is reduced with respect to the previous approach, diagnosis generation could result much faster than in OFF.

The ADD approach is conceptually similar to SIM; we consider the net obtained from the $\wedge$-fusion on the current BPN, by deleting all the sink places representing observable parameters. Given the current set of observations, we then add to N the sink places corresponding to sets P^+ and P^-. Also in this case we have to consider three distinct phases namely *observation addition*, *P-invariant computation* and *diagnosis generation* and, as in the previous case, the set of P-invariant supports we get does not contain information irrelevant to current observations.
Results concerning the comparison of the three proposed strategies are summarized in figure 3, where the average computation times of the above strategies are reported for the 48 sample cases we used. The SIM approach resulted in very high execution times, essentially because of the expensiveness of the net simplification phase. This can be see in figure 4 where cpu times of net simplification and observation addition phases are plotted. Notice also that the basic pattern of the SIM strategy in figure 3 is essentially determined by the net simplification phase. We did not investigate the possibility of directly performing the simplification on the incidence matrix of the net; our claim is that a matrix simplification will be less expensive, but it would not improve to much the result.
Between OFF and ADD, the latter strategy resulted to be better in terms of global execution time, even if without showing the huge difference of the SIM strategy. Obviously, the OFF strategy resulted in the higher computation time (with respect to both SIM and ADD strategies) for the diagnosis generation phase and for the invariant computation phase, but the fact that the latter phase is done off-line, determined the situation depicted in figure 3.

Logical and P-invariant Diagnosis Comparison

To test the performance of a P-invariant diagnostic algorithm against a classical logical approach, we chose to compare the INVADS system using the ADD strategy with an abductive diagnostic system called AID (Console *et al.* 1993). The reasons for such a direct comparison are twofold:
1. both systems rely on the same formal framework of reference we previously discussed;

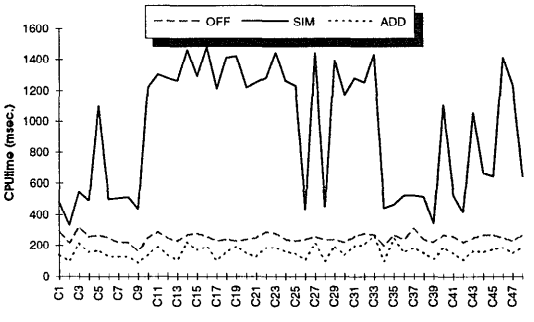

Figure 3: Comparison of different strategies for P-invariant diagnosis

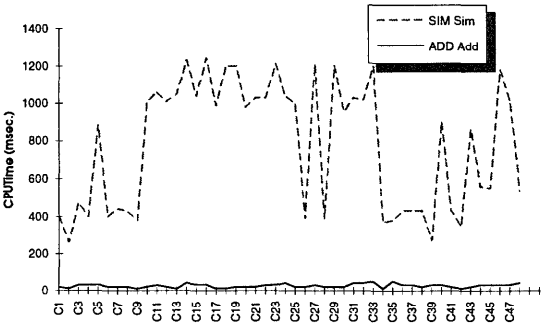

Figure 4: Net Simplification vs Observation Addition

2. both systems share the same implementation environment (a SICStus prolog implementation for SUN sparc stations).

Also for this experiment, we tested different runs of the batch of our sample cases. In particular, we measured for each case C the percentage gain of INVADS vs AID, defined as follows:

$$G(C) = \frac{T^C_{AID} - T^C_{INVADS}}{T^C_{INVADS}}$$

where T^C_{AID} and T^C_{INVADS} represent the execution times on case C of AID and INVADS respectively. Results on our car engine fault domain showed a quite good behavior of P-invariant based approach (see figure 5); the average gain resulted to be of 34.41% with some peaks of about (or even more then) 100%.

Conclusions

In the present paper, we have shown how Petri net reachability analysis could be used as a formal basis for explaining the misbehavior of a given device. We briefly discussed a net model called BPN, used to describe the behavior of the device under examination. The BPN model is not proposed as a direct tool of diagnostic knowledge representation, but rather as an analysis formalism that can be derived from other forms of knowledge representation, like for instance causal networks as described in (Portinale 1992). We concentrated on P-invariant reachability analysis, representing the starting point for the definition of an innovative approach to model-based diagnosis.

We tested the different implementation of the approach on a car engine fault domain, by getting an encouraging comparison with a classical logical approach to diagnosis. Notice also that, besides the fact that P-invariants are obtained through a linear algebra based computation (that can result more efficient than symbolic computation), parallel algorithms can be devised for this kind of approach (Marinescu, Beaven, & Stansifer 1991; Lin *et al.* 1993)). This clearly adds more interest to the net invariant approach to diagnosis, by also taking into account the fact that its complementary approach (i.e. the diagnosis based on reachability graph analysis) has been shown to be very adequate to a parallel implementation (Anglano & Portinale 1994). Future works are planned in order to compare P-invariant diagnosis also with this approach.

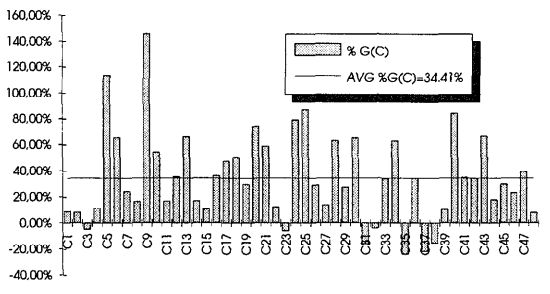

Figure 5: Percentage Gain INVADS vs AID

References

Anglano, C., and Portinale, L. 1994. B-W analysis: a backward reachability analysis for diagnostic problem solving suitable to parallel implementation. In *LNCS 815*, 39–58. Springer Verlag.

Console, L., and Torasso, P. 1991. A spectrum of logical definitions of model-based diagnosis. *Computational Intelligence* 7(3):133–141.

Console, L.; Portinale, L.; Theseider Dupré, D.; and Torasso, P. 1993. Combining heuristic and causal reasoning in diagnostic problem solving. In *Second Generation Expert Systems.* Springer Verlag. 46,68.

de Kleer, J.; Mackworth, A.; and Reiter, R. 1992. Characterizing diagnoses and systems. *Artificial Intelligence* 56(2–3):197–222.

Hamscher, W.; Console, L.; and de Kleer, J. 1992. *Readings in Model-Based Diagnosis.* Morgan Kaufmann.

Lin, C.; Chaundhury, A.; Whinston, A.; and Marinescu, D. 1993. Logical inference of Horn clauses in Petri net models. *IEEE TKDE* 5(3):416–425.

Marinescu, D.; Beaven, M.; and Stansifer, R. 1991. A parallel algorithm for computing invariants of a Petri net model. In *Proc. 4th PNPM*, 136–143.

Martinez, J., and Silva, M. 1982. A simple and fast algorithm to obtain all invariants of a generalized Petri net. In *Applications and Theory of Petri Nets.* Springer Verlag. 301–310.

Murata, T. 1989. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE* 77(4):541–580.

Peterka, G., and Murata, T. 1989. Proof procedure and answer extraction in Petri net model of logic programs. *IEEE TSE* 15(2):209–217.

Poole, D. 1989. Normality and faults in logic-based diagnosis. In *Proc. 11th IJCAI*, 1304–1310.

Portinale, L. 1992. Verification of causal models using Petri nets. *International Journal of Intelligent Systems* 7(8):715–742.

Portinale, L. 1993. *Petri net models for diagnostic knowledge representation and reasoning.* PhD Thesis, Dip. Informatica, Universita' di Torino. `ftp://ftp.di.unito.it/pub/portinal`.