Dynamic Diagnosis (Position Paper)

Chitta Baral and Tran Cao Son
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
{chitta,tson}@cs.utep.edu

Recent research in diagnosis (McIlraith, 1997; Thielscher, 1997; McIlraith, 1998) extends earlier works in diagnosis from first principles by using an action theory instead of a first-order theory to describe the correct behavior of the system in consideration. The action theory allows us to reason about actions and their effects. Thus, when an action does not yield the expected effects something did malfunction. Using regression, we can identify what is wrong initially. *Our position is that in many cases this is not enough.* The following example illustrates such a case.

Consider a simple domain where we have a light bulb connecting to a switch. The switch can be either in *on* or *off* position and can be changed by the actions *turn_on* or *turn_off*. If we turn the switch to the on position (*turn_on*) the light will be on (*light_on*) if the bulb did not burn out. If we turn the switch to the off position (*turn_off*), the light will go off ($\neg$*light_on*). Assume that the only component of the system which can be defective is the bulb and it can become defective at anytime. Initially, the light was off ($\neg$*light_on*). Executing the sequence of actions *turn_on; turn_off; turn_on*, we see the light is on, off, and off. This sequence of actions and observations can be represented by the following picture.

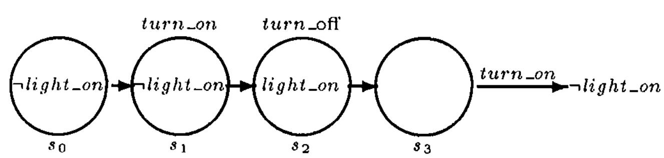

Figure 1: Does the bulb burn out ?

Intuitively, we would conclude that the bulb did burn out, i.e., $\Delta = \{bulb\}$ is a diagnosis for the system. Our argument for this conclusion is as follows: since the only breakable component is the bulb and it can be broken at anytime, the bulb must have gotten broken prior to the second time the action *turn_on* is executed (at S_3). According to the approach in (McIlraith, 1997), the whole story is inconsistent and Δ is not a diagnosis for it. This is because the approach in (McIlraith, 1997) assumes that nothing is wrong between the initial and the final situation.

Unlike the approaches in (McIlraith, 1997; Thielscher, 1997; McIlraith, 1998) we formulate diagnosis and diagnostic planning when we are given a narrative in the language $\mathcal{L}$ (Baral, Gelfond, and Provetti, 1997). A narrative is a (possibly) incomplete set of observations about the system in terms of what actions/events occured and the value of fluents at different instants in the past. The above example can be represented by a domain description D and a narrative Γ as follows.

$$D = \begin{cases} turn\_on \textbf{ causes } light\_on \textbf{ if } \neg ab(bulb) \\ turn\_off \textbf{ causes } \neg light\_on \\ ab(bulb) \textbf{ causes } \neg light\_on \end{cases}$$

$$\Gamma = \begin{cases} turn\_on \textbf{ occurs\_at } s_1 \\ turn\_off \textbf{ occurs\_at } s_2 \\ s_0 \textbf{ precedes } s_1 \\ s_1 \textbf{ precedes } s_2 \\ s_2 \textbf{ precedes } s_3 \\ \neg light\_on \textbf{ at } s_1 \\ light\_on \textbf{ at } s_2 \\ \neg light\_on \textbf{ after } turn\_on \textbf{ at } s_3 \end{cases}$$

To formulate diagnosis, we assume that the system has several components, and for each component c, there may be several fluents associated with it. One such fluent is expressed by the term $ab(c)$ denoting that the component c is broken. A system description SD is a triple $(D, OBS, COMPS)$ where

1. $COMPS = \{c_1, \ldots, c_n\}$ is a finite set of components;
2. D is a set of laws (both static and dynamic) expressing the behavior of the system and in addition contains dynamic laws of the form $break(c)$ **causes** $ab(c)$, for each component c in $COMPS$.
3. OBS is a collection of observations starting from the situation s_1 (i.e., fluent facts, occurrence facts, and precedence facts are only about situations s_1 and beyond and are not about s_0) and in addition contains the precedence fact s_0 **precedes** s_1.

Here, the action $break(c)$ is used to explain unexpected observations about $ab(c)$. By requiring that observations start from the situation s_1, we are able to express that everything is fine in the initial situation s_0. This is expressed by the set $OBS_0 = \{\neg ab(c) \ \textbf{at} \ s_0 \mid c \in COMPS\}$. Thus, a system description $(D, OBS, COMP)$ needs a diagnosis if $(D \setminus D_{ab}, OBS \cup OBS_0)$ does not have a model where D_{ab} denotes the set of causal laws whose effect is $ab(c)$ for some $c \in COMPS$. A diagnosis of SD w.r.t. a state s is then defined by the set Δ, $\Delta \subseteq COMPS$, such that there exists a model M of $(D, OBS \cup OBS_0)$ and $M \models ab(c) \ \textbf{at} \ s$ for every $c \in \Delta$.

We prove that our approach, when used in system descriptions without actions, is sound with respect to diagnosis from first principle. We also give sufficient conditions for its completeness with respect to diagnosis from first principle.

We define a diagnostic plan as a (possible) conditional plan which, when executed from the current state, will identify the "real" diagnosis from the many "hypothetical" diagnoses of a system. This differs from traditional testing where only one action is considered at a time. We show that discriminating tests, as it is considered in the literature (McIlraith, 1994), can be viewed as a special case of diagnostic plans under certain conditions.

References

C. Baral, M. Gelfond, and A. Provetti: Representing Actions: Laws, Observations, and Hypotheses, Journal of LP (31) 1–3, 1997, pp. 201–243.

S. A. McIlraith. Generating Tests using Abduction In Proc. of KR 94, pp. 449–460.

S. A. McIlraith: Representing Actions and State Constraints in Model-Based Diagnosis, AAAI'97, pp. 43-49.

S. A. McIlraith: Explanatory Diagnosis: Conjecturing actions to explain observations, KR 98.

M. Thielscher. Diagnosis in Dynamic Domains. In ETAI, Vol. 2(1997): nr 11.