

Modellbasierte Diagnosesysteme

Diagnose: Identifikation eines vorliegenden Fehlers (Krankheit) auf der Basis von Beobachtungen (Symptomen) und Hintergrundwissen über das System

2 Arten von Diagnosesystemen: abduktiv vs. konsistenzbasiert

Typ I: abduktive Diagnose

Gegeben: Information über

- 1) Wirkung von Fehlern H (Implikationen)
- 2) beobachtete Symptome S (Atome)
- 3) nicht aufgetretene Symptome N (negierte Literale)

Gesucht: Erklärung für S, d.h.:

minimale Menge von Fehlern F, so dass $F \cup H \models S$ und $F \cup H \cup N$ konsistent.

Deduktion: Ableiten von implizitem Wissen

Alle A sind B

c ist ein A

c ist ein B

Abduktion: Finden von Erklärungen für Beobachtungen

Alle A sind B

c ist ein B

c ist ein A

Induktion: Ableiten allgemeiner Regel aus bisherigen Beobachtungen

c ist ein A (c_1 ist ein A, ..., c_n ist ein A)

c ist ein B (c_1 ist ein B, ..., c_n ist ein B)

Alle A sind B

Beispiel:

H: Masern -> Fieber & Ausschlag
Migräne -> Kopfschmerzen & Übelkeit
Grippe -> Kopfschmerzen & Gliederschmerzen & Fieber

N: leer

O: Fieber => 2 Erklärungen: Masern, Grippe
Fieber, Ausschlag => 1 Erklärung: Masern
Kopfschmerzen => 2 Erklärungen: Migräne, Grippe
Kopfschmerzen, Fieber => 1 Erklärung: Grippe

N: \neg Übelkeit

O: Kopfschmerzen => 1 Erklärung: Grippe
Kopfschmerzen, Ausschlag => 1 Erklärung: Grippe & Masern

Typ II: konsistenzbasierte Diagnose:

Die abduktive Diagnose funktioniert immer dann, wenn man ausreichende Information darüber hat, zu welchen Effekten (Symptomen) Fehler führen. Oft verfügt man aber, z.B. bei technischen Systemen, über genaue Information über das korrekte Verhalten. In diesem Fall ist die konsistenzbasierte Diagnose vorteilhaft. Hier geht man von der Annahme aus, dass Komponenten korrekt arbeiten – außer dies führt zu Widersprüchen mit Beobachtungen.

Gegeben: Menge K von Komponenten sowie Information über

- 1) Wirkung der Komponenten im fehlerfreien Fall (Modell M)
- 2) Beobachtungen O

Gesucht: minimale Menge K' von Komponenten, so dass die Annahme der Korrektheit aller Komponenten in $K \setminus K'$ konsistent mit $M \cup O$ ist.

Teilaufgaben bei der konsistenzbasierten Diagnose:

- a) Erkennen von Abnormalitäten (Konflikten)
- b) Generieren und Testen von Hypothesen (Kandidaten)
- c) Diskriminieren zwischen Hypothesen (Meßpunktauswahl)

Prototypisches Beispiel **GDE** (General Diagnostic Engine), de Kleer, Williams

modelliert Struktur von Schaltkreisen, um deren Verhalten vorherzusagen
ermittelt durch Constraint-Propagierung erwartete Werte,
Nichtübereinstimmung mit gemessenen Werten initiiert Diagnose

a) Konflikterkennung

Ist System in Zustand, der Diagnose erfordert?

Konflikt: Menge von Komponenten K , so dass die Annahme des korrekten Verhaltens aller Elemente in K zur Prognose eines Wertes führt, die im Widerspruch zu Beobachtung steht.

System überprüft Prognosen aus Annahmemenge A erst, wenn alle Teilmengen von A konfliktfrei sind. Nur minimale Konflikte werden so erzeugt.

Zur Konflikterkennung verwendet **GDE** ein **ATMS** (assumption based truth maintenance system):

ATMS:

input: aussagenlogische Hornklausen J (justifications):

(Instanzen von Regeln, die das Diagnosesystem angewendet hat)

Menge von Annahmen A (assumptions):

(Komponenten des zugrundeliegenden Systems, die als ok angenommen werden)

output: für jede Proposition Mengen von Annahmen, unter denen sie abgeleitet werden kann.

interagiert mit Problemlöser:

Problemlöser:

- verfügt über Modell des jeweiligen Systems, das das Verhalten seiner Komponenten (im korrekten und falls möglich auch im inkorrekten Fall) beschreibt
- erfragt von Benutzer Daten
- wendet Regeln an, die auf der Basis von Annahmen über die Korrektheit von Komponenten neue Werte aus den Daten ermitteln
- übergibt entsprechende justifications an das ATMS
- lässt sich vom ATMS minimale Konflikte berechnen und erzeugt daraus Hypothesen

ATMS

- verwaltet Abhängigkeitsgraphen auf der Basis der vom Problemlöser gelieferten justifications
- berechnet Labels, das Label des Knotens **false** enthält alle minimalen Konflikte

Verwalteter Abhängigkeitsgraph:

Knoten: Propositionen

Kanten: justifications

Kontradiktion: ausgezeichnete Knoten **false**.

Label des Knotens k: Menge L von Mengen von Annahmen E_i , so dass aus $E_i \cup J \vdash k$

Annahmen: Knoten k mit Label $\{\{k\}\}$

Elemente des labels von **false** heißen auch nogoods.

Bedingungen für Labels: sei n Knoten, $\{E_1, \dots, E_n\}$ sein Label

1. Korrektheit: $E_i \cup J \vdash n$.
2. Konsistenz: nicht $E_i \cup J \vdash \text{false}$ (falls n \neq **false**)
3. Vollständigkeit: wenn $E \cup J \vdash n$ dann gibt es E_i mit E_i Teilmenge von E
4. Minimalität: es gibt keine E_i, E_j im Label, so dass E_i echte Teilmenge von E_j

ATMS update: gegeben Abhängigkeitsgraph mit korrekten Labels, Problemlöser liefert neue justification $n_1 \wedge \dots \wedge n_k \Rightarrow n$

1. berechne neue tentative Labels für n (alle möglichen Kombinationen der Labels der n_j).
2. eliminiere nogoods und subsumierte (nichtminimale) Annahmemengen.
3. propagiere ggfs. neues Label von n im Netz.

Beispiel:

e: $\{\{A,B\},\{C\}\}$

f: $\{\{A\},\{D\}\}$

nogood $\{C,D\}$

e ist wahr, falls A,B gilt oder falls C gilt

f ist wahr, falls A gilt oder falls D gilt

C und D können nicht beide wahr sein

neue justification: $e \wedge f \Rightarrow g$

1. Schritt: $g: \{\{A,B\},\{A,B,D\},\{C, A\},\{C, D\}\}$

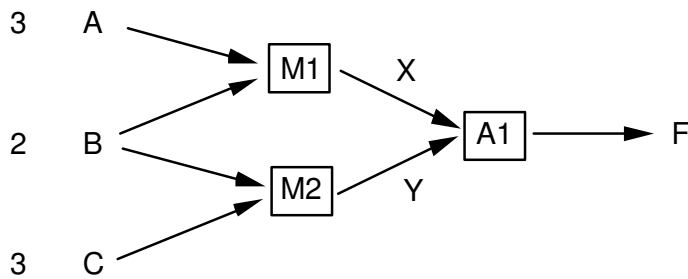
2. Schritt: $g: \{\{A,B\},\{C, A\}\}$

alle Kombinationen der Labels von e, f

daraus nicht-minimale und nogoods weg

Label von **false** enthält gerade alle minimalen Konfliktmengen!

Beispiel: System aus 2 Multiplizierern M1 und M2 und einem Addierer A1:

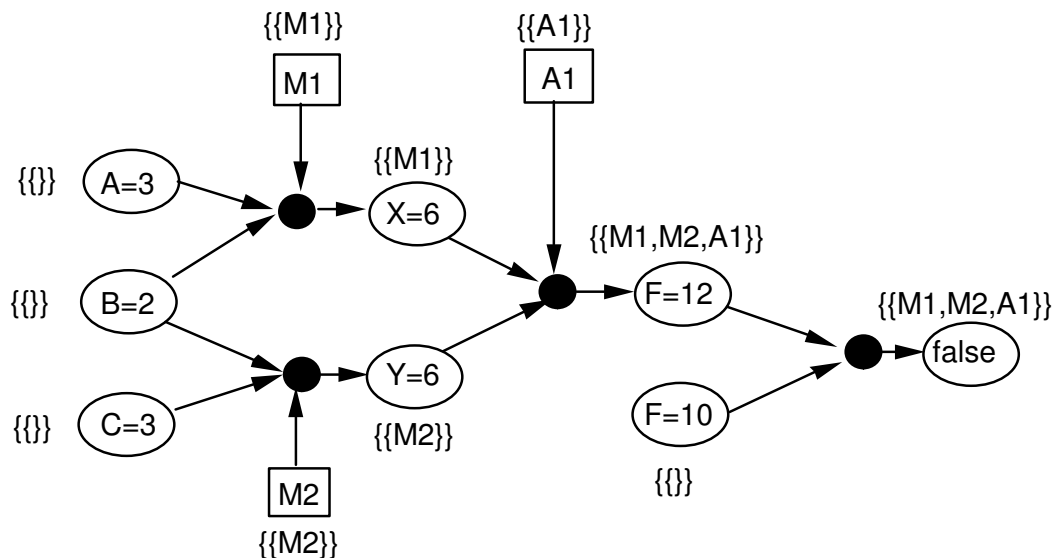


Der Problemlöser verfügt über die Beschreibung des Systems und gemessene Daten $A=3$, $B=2$, $C=3$, $F=10$.

Er propagiert Werte auf Basis der verfügbaren Komponentenbeschreibungen und übergibt jeweils entsprechende justifications an das ATMS. Dabei wird folgende Situation erreicht:

Assumptions:	M1, M2, A1	bedeutet: M1 ok, M2 ok, A1 ok
justifications:	$\Rightarrow A=3$	Beobachtungen
	$\Rightarrow B=2$	
	$\Rightarrow C=3$	
	$\Rightarrow F=10$	
	$A=3 \wedge B=2 \wedge M1 \Rightarrow X=6$	propagierter Wert, falls M1 ok
	$B=2 \wedge C=3 \wedge M2 \Rightarrow Y=6$	propagierter Wert, falls M2 ok
	$X=6 \wedge Y=6 \wedge A1 \Rightarrow F=12$	propagierter Wert, falls A1 ok
	$F=12 \wedge F=10 \Rightarrow \text{false}$	Widerspruch entdeckt

Mit diesen Annahmen und justifications berechnet das ATMS folgende Labels



Label von false: $\{M1, M2, A1\}$.

Damit ist ein Konflikt entdeckt: mindestens eine dieser 3 Komponenten muss defekt sein.

b) Hypothesengenerierung

Auf Basis der durch das ATMS berechneten Konflikte werden Hypothesen (mögliche Diagnosen) ermittelt. Dabei geht man davon aus, dass nur minimale Mengen von (defekten) Komponenten von Interesse sind, da Komponenten im Normalfall funktionieren. Gleichzeitig muss aber jeder Konflikt berücksichtigt werden, d.h. aus jedem Konflikt muss mindestens 1 Komponente in einer Hypothese enthalten sein. Das führt zu folgender Definition:

Def.: Eine Hypothese ist eine minimale Menge H von Komponenten, so dass $H \cap K \neq \emptyset$ für alle Konflikte K (minimal covering set).

Wird neue Konfliktmenge K_1 gefunden, so dass $H \cap K_1 = \emptyset$, so wird (für alle k aus K_1) $H \cup \{k\}$ anstelle von H als neue Hypothese verwendet.

Generell werden nur minimale (bzgl. der Teilmengenrelation) Hypothesen erzeugt. Testen von Hypothesen: wenn Fehlermodelle existieren, dann erzeugen Hypothesen erwartete Werte, sind widerlegt, wenn erwartete Werte nicht beobachtet werden.

Hauptproblem: riesige Anzahl möglicher Hypothesen.

single fault (Hyp. enthalten genau 1 Element) vs. multiple fault assumption:

- single fault assumption löst Problem, aber oft unrealistisch: defekte Komponente kann zu anderen Defekten führen, Krankheit kann Patienten schwächen und anfällig für andere Krankheiten machen, ...
- multiple faults: erhebliches kombinatorisches Problem:
bei nur 1 Fehlermode 2^n mögliche Fehler (n Anzahl der Komponenten)
50 Komponenten, bereits bei maximal 3 defekten Komponenten ca 20K mögliche Hypothesen.

deshalb häufig Hierarchien und Ranking von Hypothesen aufgrund von Wahrscheinlichkeiten

c) Hypothesendiskriminierung: Bestimmung der geeignetsten Messpunkte

Hier werden Methoden aus der Informationstheorie verwendet, um Messungen mit dem größten Informationsgewinn zu ermitteln.