Konštrukcia diagnózy chybného správania sa zložitých systémov

Vladimír Kvasnička
Ústav aplikovanej informatiky FIIT STU

Abstrakt

Predpokladajme, že máme ***systém***, ktorý obsahuje navzájom poprepájané súčiastky – komponenty, pričom pozorujeme také ***správanie sa systému***, ktoré je konfliktné (nie je konzistentné) s tým, ako sa má správať. ***Diagnostický problém*** spočíva v určení tých súčiastok – komponent, ktorých nekorektné (abnormálne) správanie sa vysvetľuje diskrepanciu medzi pozorovaným a korektným správaním sa systému.

Literatúra

- Raymond Reiter: A Theory of Diagnosis from First Principles. *Artificial Intelligence* **32** (1987), pp. 57-95
- Johan de Kleer, J., B.C. Williams, Diagnosing multiple faults, *Artificial Intelligence* **32** (1987) pp. 97-130

Raymond Reiter (1939-2002)

Johan de Kleer

A Theory of Diagnosis from First Principles

Raymond Reiter
Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 1A4; The Canadian Institute for Advanced Research

Recommended by Johan de Kleer and Daniel G. Bobrow

ABSTRACT
Suppose one is given a description of a system, together with an observation of the system's behaviour which conflicts with the way the system is meant to behave. The diagnostic problem is to determine those components of the system which, when assumed to be functioning abnormally, will explain the discrepancy between the observed and correct system behaviour.

We propose a general theory for this problem. The theory requires only that the system be described in a suitable logic. Moreover, there are many such suitable logics, e.g. first-order, temporal, dynamic, etc. As a result, the theory accommodates diagnostic reasoning in a wide variety of practical settings, including digital and analogue circuits, medicine, and database updates. The theory leads to an algorithm for computing all diagnoses, and to various results concerning principles of measurement for discriminating among competing diagnoses. Finally, the theory reveals close connections between diagnostic reasoning and nonmonotonic reasoning.

1. Introduction

In the theory and design of diagnostic reasoning systems there appear to be two quite different approaches in the literature.

In the first approach, often referred to as diagnosis from first principles, one begins with a description of some system—a physical device or real world setting of interest, say—together with an observation of the system's behaviour. If this observation conflicts with the way the system is meant to behave, one is confronted with a diagnostic problem, namely, to determine those system components which, when assumed to be functioning abnormally, will explain the discrepancy between the observed and correct system behaviour. For solving this diagnostic problem from first principles, the only available information is the system description, i.e. its design or structure, together with the observation(s) of the system behaviour. In particular, no

Artificial Intelligence **32** (1987) 57–95
0004-3702/87/$3.50

Diagnosing Multiple Faults

Johan de Kleer
Intelligent Systems Laboratory, XEROX Palo Alto Research Center, Palo Alto, CA 94304, U.S.A.

Brian C. Williams
Artificial Intelligence Laboratory, MIT, Cambridge, MA 02139, U.S.A.

Recommended by Judea Pearl

ABSTRACT
Diagnostic tasks require determining the differences between a model of an artifact and the artifact itself. The differences between the manifested behavior of the artifact and the predicted behavior of the model guide the search for the differences between the artifact and its model. The diagnostic procedure presented in this paper is model-based, inferring the behavior of the composite device from knowledge of the structure and function of the individual components comprising the device. The system (GDE—general diagnostic engine) has been implemented and tested on many examples in the domain of troubleshooting digital circuits.

This research makes several novel contributions: First, the system diagnoses failures due to multiple faults. Second, failure candidates are represented and manipulated in terms of minimal sets of violated assumptions, resulting in an efficient diagnostic procedure. Third, the diagnostic procedure is incremental, exploiting the iterative nature of diagnosis. Fourth, a clear separation is drawn between diagnosis and behavior prediction, resulting in a domain (and inference procedure) independent diagnostic procedure. Fifth, GDE combines model-based prediction with sequential diagnosis to propose measurements to localize the faults. The normally required conditional probabilities are computed from the structure of the device and models of its components. This capability results from a novel way of incorporating probabilities and information theory into the context mechanism provided by assumption-based truth maintenance.

1. Introduction

Engineers and scientists constantly strive to understand the differences between physical systems and their models. Engineers troubleshoot mechanical systems or electrical circuits to find broken parts. Scientists successively refine a model based on empirical data during the process of theory formation. Many

Artificial Intelligence **32** (1987) 97–130
0004-3702/87/$3.50

ftp://math.chtf.stuba.sk/pub/vlado/Diagnosis/reiter-diagnosis%20AIJ%20reprint.pdf

ftp://math.chtf.stuba.sk/pub/vlado/Diagnosis/deKleer_Diagnosing%20Multiple%20Faults%20AIJ%20reprint.pdf

Priesvitka 3

Úvodné poznámky

- Stanovenie diagnózy zložitých systémov (v medicíne, vo veľkých výrobných zariadeniach, elektronických obvodov, a pod.) patrí od vzniku umelej inteligencie pred polstoročím k jedným z jej základných problémov.

- V 60. a 70. rokoch minulého storočia boli vytvorené expertné systémy pre stanovenie diagnózy v medicíne, ktoré pokrývajú rôzne oblasti internej medicíny, psychiatrie, farmakológie a pod. (pozri učebnicu expertných systémov od Poppera a Kelemena [xx]).

- Základnou paradigmou týchto prístupov je ***model*** systému, ktorý je vytvorený pomocou súboru poznatkov – pravidiel.

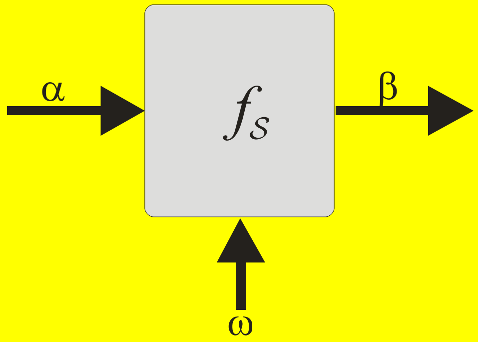
- V jazyku teórie systémov, systém je špecifikovaný jeho elementmi a interakciami medzi nimi.

- Konštrukcia diagnózy pre interpretáciu chybného správanie sa systému spočíva v hľadaní takých chybných elementov a spojov medzi nimi (t. j. poznatkov o danom a ich vzájomnej interakcii), ktoré sú schopné vysvetliť pozorované chybné správanie sa systému.

- Hľadanie týchto vhodných poznatkov, ktoré nám interpretujú chybné správanie sa systému, je ***obvykle*** založené na inferenčnom aparáte klasickej výrokovej (a predikátovej) logiky.

- Cieľom tejto prednášky je ukázať, že klasický inferenčno-logický prístup k riešeniu problému diagnózy môže byť nahradený iným ***numerickým princípom minimalizacie účelovej funkcie***, ktorá nám popisuje „vzdialenosť" medzi pozorovaným a očakávaným (teoretickým) správaním sa systému v závislosti od korektnosti alebo nekorektnosti (chybovosti) podsystémov študovaného systému.

Budeme študovať konštrukciu diagnózy ***binárneho systému*** $\mathcal{S}$



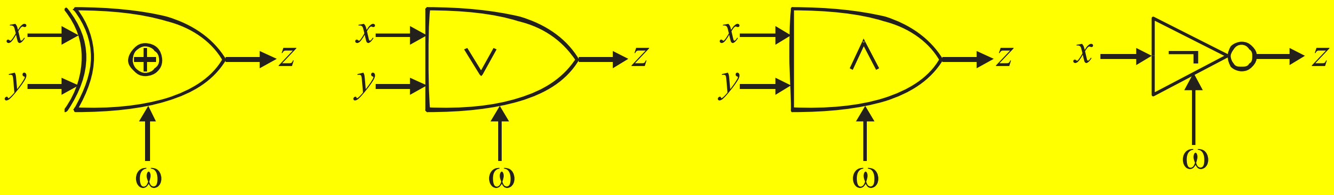
$$f_{\mathcal{S}}(\boldsymbol{\omega}):\{0,1\}^m \to \{0,1\}^n$$
$$\boldsymbol{\beta} = f_{\mathcal{S}}(\boldsymbol{\alpha};\boldsymbol{\omega})$$

$\boldsymbol{\omega} \in \{0,1\}^p$ je „parameter" zobrazenia $f_{\mathcal{S}}(\boldsymbol{\omega})$,

$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \{0,1\}^m$ nazývame ***vstup*** binárneho systému

$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m) \in \{0,1\}^n$ sa nazýva ***výstup*** binárneho systému

Z teórie Boolových funkcií vieme, že každá Boolova funkcia môže byť vyjadrená pomocou elementárnych (binárnych alebo unárnych) Boolových funkcií, ktoré sú priradené logickým spojkám XOR, OR, AND a NOT,



A B C D

Elementárne Boolove funkcie sú špecifikované pomocou logických spojok a parametru ω

$$z = XOR\left(x, y; \omega\right) = \left(\left(x \oplus y\right) \wedge \omega\right) \vee \left(\left(x \equiv y\right) \wedge \neg\omega\right)$$

$$z = OR\left(x, y; \omega\right) = \left(\left(x \vee y\right) \wedge \omega\right) \vee \left(\left(\neg x \wedge \neg y\right) \wedge \neg\omega\right)$$

$$z = AND\left(x, y; \omega\right) = \left(\left(x \wedge y\right) \wedge \omega\right) \vee \left(\left(\neg x \vee \neg y\right) \wedge \neg\omega\right)$$

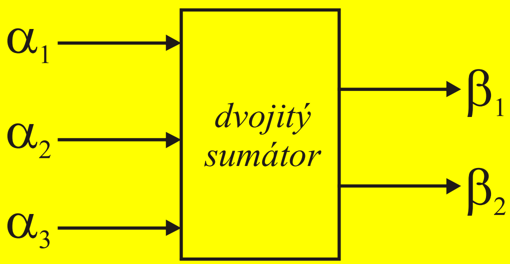
$$z = NOT\left(x; \omega\right) = \left(\neg x \wedge \omega\right) \vee \left(x \wedge \neg\omega\right)$$

Dvojitý sumátor

Ako ilustračný príklad binárneho systému $\mathcal{S}$ budeme študovať ***dvojitý sumátor*** (ang. *full adder*)

$$\alpha_1 + \alpha_2 + \alpha_3 = \beta_1\beta_2$$

K takto špecifikovanému sumátoru priradíme tabuľku vstupných a výstupných hodnôt

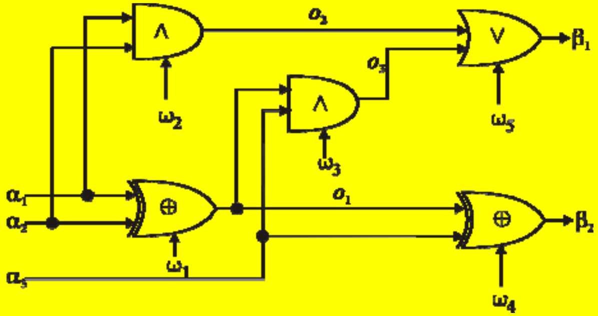


	vstup			*výstup*	
#	α_1	α_2	α_1	β_1	β_2
1	0	0	0	0	0
2	0	0	1	0	1
3	0	1	0	0	1
4	0	1	1	1	0
5	1	0	0	0	1
6	1	0	1	1	0
7	1	1	0	1	0
8	1	1	1	1	1

Použitím všeobecnej teórie syntézy Boolových funkciízadaných tabuľkou ich funkčných hodnôt, dostaneme tieto výrazy pre Boolovu funkciu, ktorá reprodukuje tabuľku z

$$\beta_1 = \bar{\alpha}_1\alpha_2\alpha_3 + \alpha_1\bar{\alpha}_2\alpha_3 + \alpha_1\alpha_2\bar{\alpha}_3 + \alpha_1\alpha_2\alpha_3 = \left(\alpha_1 \oplus \alpha_2\right)\alpha_3 + \alpha_1\alpha_2$$

$$\beta_2 = \bar{\alpha}_1\bar{\alpha}_2\alpha_3 + \bar{\alpha}_1\alpha_2\bar{\alpha}_3 + \alpha_1\bar{\alpha}_2\bar{\alpha}_3 + \alpha_1\alpha_2\alpha_3 = \alpha_1 \oplus \alpha_2 \oplus \alpha_3$$

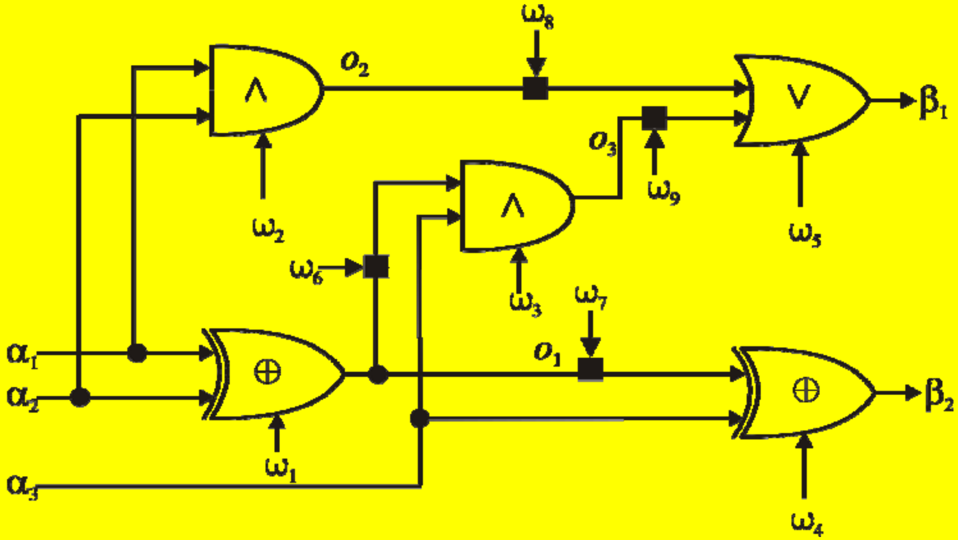



```
procedure full_adder(α1,α2,α3: Boolean; {input_activities}
                          var β1,β2: Boolean; {output_activities}
                          ω1, ω2, ω3, ω4, ω5: Boolean {parameters});
var o1,o2,o3 : Boolean; {inner activities}
begin o1:=XOR(α1,α2, ω1); o2:=AND(α1,α2, ω2); o3:=AND(α3,o1, ω3);
|     β1:=OR(o3,o2, ω5); β2:=XOR(o1,α3, ω4);
end;
```

- Binárny vektor parametrov $\boldsymbol{\omega}=(\omega_1,\omega_2,...,\omega_5)$ slúži na špecifikáciu tých elementárnych Boolovych brán ktoré su korektné (ω_i = 1) resp. nekorektné (ω_i = 0). V prípade, že všetky jeho komponenty sú jednotkové, potom binárny obvod 16.7 nevykazuje chybu, všetky elementárne Boolove brány pracujú korektne.

- V použitom prístupe chybovosť binárneho systému $\mathcal{S}$ spočíva len v možnosti chybnej aktivity niektorých jeho elementárnych brán a nie na prerušení spoja medzi dvoma elementárnymi bránami (t. j. postulujeme, že spoje v danom binárnom systéme $\mathcal{S}$ sú *a-priori* korektné). Pri interpretácii chybného správania daného binárneho systému sústreďujeme pozornosť len na možnú nekorektnosť elementárnych brán a nie na nekorektnosť spojov medzi nimi.

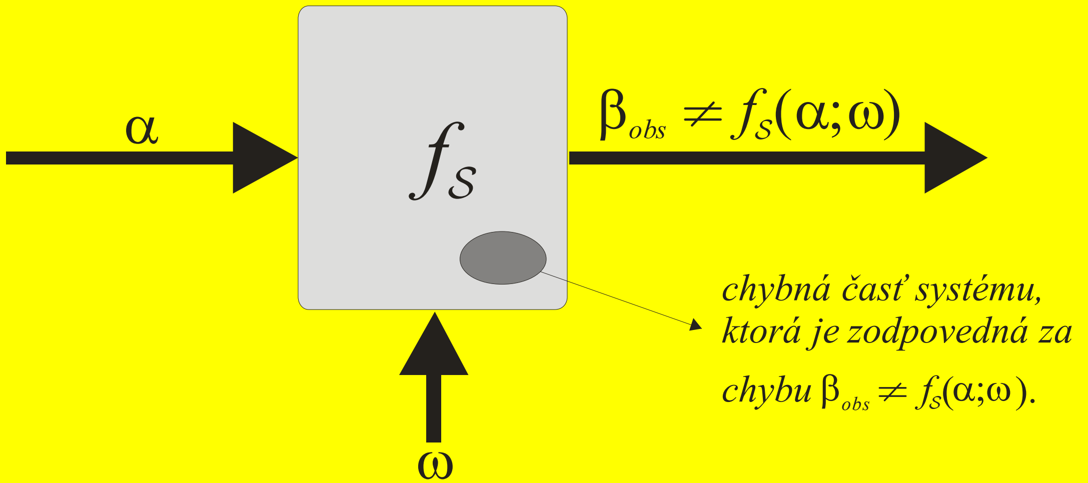
Všeobecnosť použitého prístupu k stanoveniu diagnózy binárneho systému $\mathcal{S}$ môžeme zvýšiť tak, že zavedieme ďalšie binárne parametre $\left(\omega_{p+1},\omega_{p+2},\ldots,\omega_{p+q}\right)$, kde ω_i špecifikuje, či i-tý spoj v binárnom systéme je neprerušený ($\omega_i = 1$) alebo prerušený ($\omega_i = 0$).



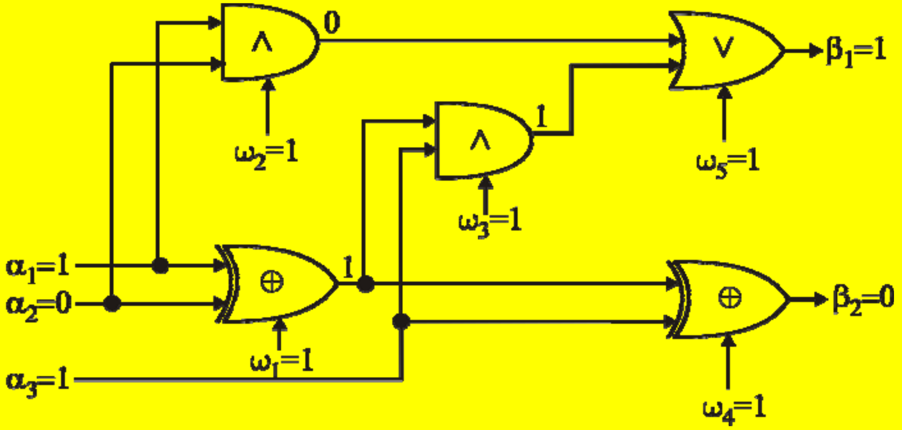
Formálna špecifikácia systému

Ak poznáme „matematický model“ systému (kde korektnosť jednotlivých elementárnych brán je špecifikovaná binárnym vektorom ω), potom sme schopní spočítať jeho odozvu $f_{\mathcal{S}}(\alpha;\omega)$ na ľubovoľný vstup α, a teda aj rozhodnúť, či nejaký aktuálny výstup systému β je v súhlase s očakávaním výstupom $f_{\mathcal{S}}(\alpha;\omega)$, ktorý je reakciou na vstup α.

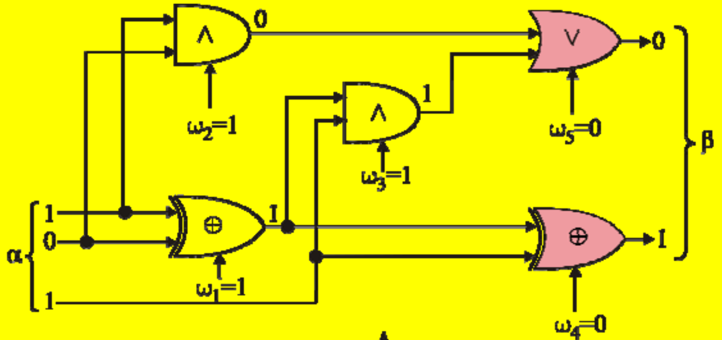
V prípade, že platí $f_{\mathcal{S}}(\alpha;\omega) \neq \beta$, t. j. očakávaný výstup sa líši od skutočného výstupu, potom vystupuje do popredia otázka „***prečo***“. Toto je problém stanovenia diagnózy, t. j. určenia tej (alebo tých) elementárnej Boolovej funkcie podieľajúcej sa na konštrukcii celkovej modelovej funkcii $f_{\mathcal{S}}$, ktorá je zodpovedná (alebo ktoré sú zodpovedné) za tento chybný výsledok,



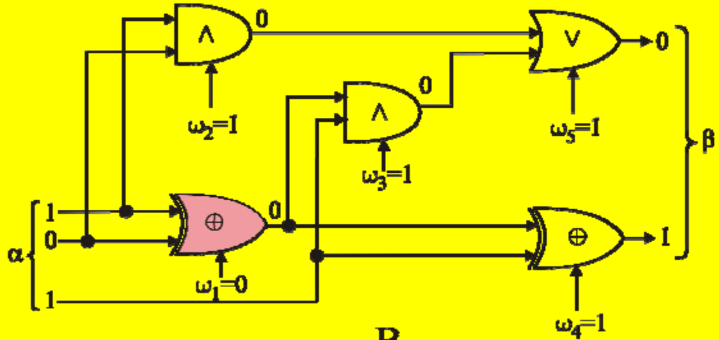
Znázornenie pozorovania nad binárnym systémom, ktorý je reprezentovaný Boolovou funkciou f. Pod pozorovaním rozumieme zistenie nesúhlasu medzi pozorovaným výstupom β_{obs} a očakávanou odozvou $f_{\mathcal{S}}(\alpha)$ na vstup α, t. j. $\beta_{obs} \neq f_{\mathcal{S}}(\alpha;\, \omega)$. Táto experimentálna skutočnosť je vysvetlená takou diagnózou, ktorá pomocou parametrov ω_1, ω_2, ..., ω_p stanoví takú časť systému (reprezentovanú elementárnymi Boolovymi funkciami), ktorá vysvetlí pozorovanie.



Znázornenie dvojitého sumátora z obrázku 16.5, kde vstup je $\boldsymbol{\alpha} = (\alpha_1/1, \alpha_2/0, \alpha_3/1)$ a výstup je $\boldsymbol{\beta} = (\beta_1/1, \beta_2/0)$, pričom $f_{\mathcal{S}}(\boldsymbol{\alpha};\boldsymbol{\omega}) = \boldsymbol{\beta}$ (korektný výstup), pre $\boldsymbol{\omega} = (\omega_1/1, \omega_2/1, \omega_3/1, \omega_4/1, \omega_5/1)$.

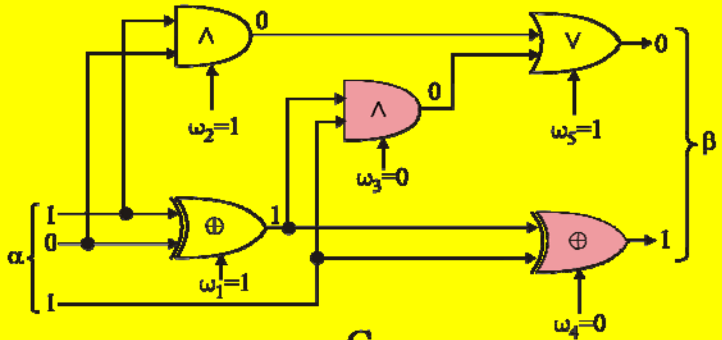


A



B

Chybny výstup
s identifikáciou
chybnych súčiastok



C

Priesvitka 14

Ako špecifikovať chybné elementárne bloky?

Konštrukcia optimálnej diagnózy je problém systematického prehľadávania nad potenčnou množinou

$$\mathcal{P}\left(\{1,2,\ldots,p\}\right)=\left\{\Delta;\Delta\subseteq\{1,2,..,p\}\right\}$$

Pre danú podmnožinu $\Delta\in\mathcal{P}\left(\{1,2,..,p\}\right)$, parametre ω_i $(i\in\Delta)$ sú špecifikované takto:

$$i\in\Delta\Rightarrow\omega_i=0$$

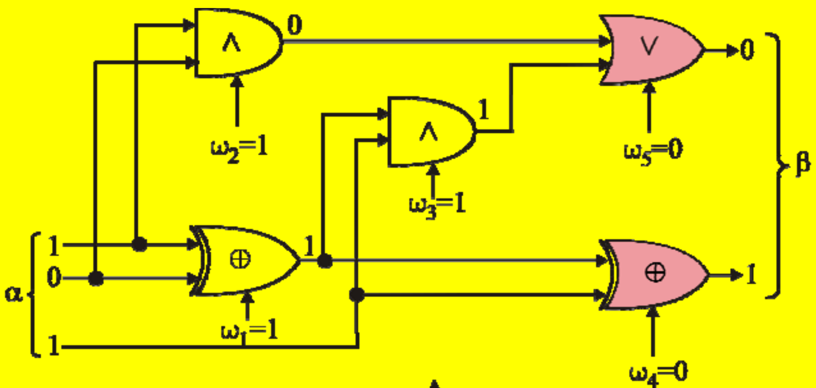
$$i\notin\Delta\Rightarrow\omega_i=1$$

Definícia.

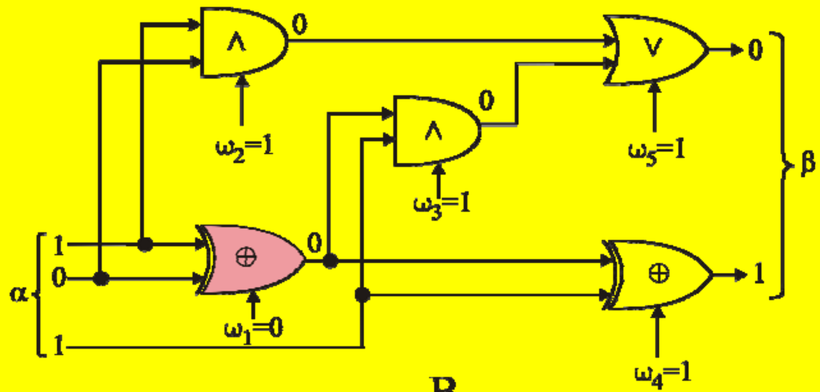
(1) Množina $\Delta \in \mathcal{P}\left(\{1,2,..,p\}\right)$ sa nazýva ***vhodná*** pre interpretáciu binárneho systému $\mathcal{S}$ s danou dvojicou vstupno-výstupných binárnych vektorov $\boldsymbol{\alpha} \in \{0,1\}^m$ a $\boldsymbol{\beta} \in \{0,1\}^n$ práve vtedy, ak vypočítaný vektor $\tilde{\boldsymbol{\beta}} = f_{\mathcal{S}}\left(\boldsymbol{\alpha},\boldsymbol{\omega}\right)$ je totožný s požadovaným výstupným vektorom $\boldsymbol{\beta}$, t. j. $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}$.

(2) Vhodná množina Δ (vzhľadom k dvojici vstupno-výstupných vektorov $\boldsymbol{\alpha}$ a $\boldsymbol{\beta}$) sa nazýva ***minimálna*** práve vtedy, ak nemôžeme z nej odstrániť ani jeden element tak, aby sme nestratili jej vlastnosť vhodnosti.

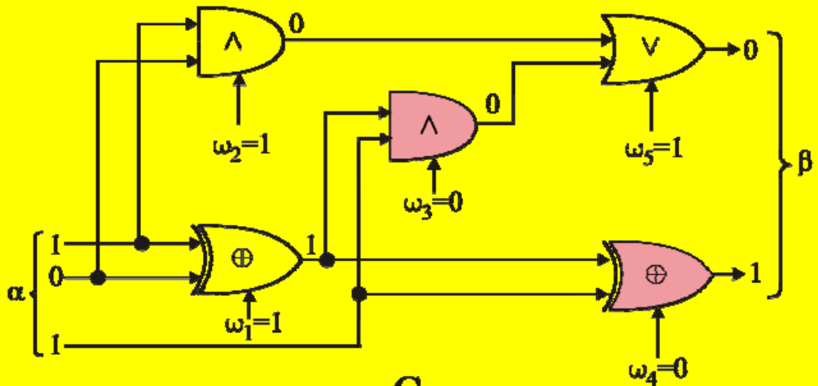
Poznámka: Podmienka minimálnosti zabezpečuje takú vlastnosť ľubovoľnej dvojice množín $\Delta' \subseteq \Delta$, že ak množina Δ je minimálna a vhodná, potom podmnožina Δ' nemôže byť vhodnou množinou.



A



B



C

$$\Delta_1 = \{4,5\}, \Delta_2 = \{1\}, \Delta_3 = \{3,4\}$$

Heuristika pre stanovenie diagnózy

- Nech „tréningová “množina
  - $\mathcal{A}_{train} = \{\alpha/\beta_{obs}\}$

  obsahuje dvojice aktivít, kde β_{obs} je „pozorovaná“ (alebo „experimentálna“) odozva binárneho systému $\mathcal{S}$ na vstupné aktivity α, pričom požadované aktivity (vypočítané pomocou modelu binárneho systému) sú $\beta = f_{\mathcal{S}}(\alpha;\omega)$.
- K takto definovanej množine vstupno-výstupných aktivít definujme účelovú funkciu

$$E(\omega) = \sum_{\alpha/\beta_{obs}} \left|\beta_{obs} - f_{\mathcal{S}}(\alpha;\omega)\right|$$

- Naším cieľom bude stanoviť takú hodnotu parametrov ω_{opt}, ktorá minimalizuje (na nulovú hodnotu) funkcionál $E(\omega)$

$$\omega_{opt} = arg \min_{\omega\in\{0,1\}^p} E(\omega)$$

- Vektory $\boldsymbol{\omega}$ môžeme jednoznačne prepísať do tvaru podmnožín $\Delta \in \mathcal{P}\left(\{1,2,...,p\}\right)$ patriacich do potenčnej množiny vytvorenej pre číselnú množinu $\{1,2,...,p\}$ obsahujúcej indexy elementárnych logických brán, potom

$$\Delta_{opt} = arg\min_{\Delta \in \mathcal{P}} E\left(\Delta\right)$$

Veta.
Majme dve množiny $\Delta' \subset \Delta$, ak Δ' je vhodná minimálna množina, potom množina Δ nie je vhodnou minimálnou množinou.

Pomocou tejto vety môže podstatne zredukovať strom riešení pri generovaní účelovej funkcie E pre podmnožiny Δ. Postupne budeme zostrojovať množinu riešení $\mathcal{D}=\left\{\Delta_{opt}^{(1)},\Delta_{opt}^{(2)},...\right\}$, pričom jej elementy vyhovujú podmienkam

$$\left(\forall\Delta_{opt}\in\mathcal{D}\right)\left(E\left(\Delta_{opt}\right)=0\right)$$

$$\left(\forall\Delta_{opt},\Delta'_{opt}\right)\left(\Delta_{opt}\not\subseteq\Delta'_{opt}\right)$$

Kde prvá podmienka hovorí, že každý element množiny $\mathcal{D}$ je riešením optimalizačného problému a druhá podmienka hovorí, že medzi riešeniami z $\mathcal{D}$ neexistuje vzťah inklúzie.

Heuristika (stratégia riešenia optimalizačného problému)

1. krok. Riešenie opt. problému hľadáme v tvare jednoprvkovej množiny $\Delta = \{i\}$, kde $i \in \{1, 2, ..., p\}$ je index elementárnej logickej brány. Množina riešení $\mathcal{D}$ obsahuje všetky tieto elementárne riešenia s jednotkovou kardinalitou.

2. krok. Riešenie opt. Problému hľadáme v tvare dvojprvkovej množiny $\Delta = \{i < j\}$, kde $i, j \in \{1, 2, ..., p\}$. Toto riešenie Δ_{opt} je zahrnuté do množiny $\mathcal{D}$ práve vtedy, ak platí predchádzajúca veta, t. j. táto množina $\mathcal{D}$ neobsahuje také riešenie Δ', ktoré by bolo podmnožinou aktuálneho riešenia $\Delta = \{i, j\}$.

$$\left(\Delta'_{opt} \in \mathcal{D}\right)\left(\Delta'_{opt} \not\subseteq \Delta_{opt}\right) \Rightarrow \mathcal{D} := \mathcal{D} \cup \left\{\Delta_{opt}\right\}$$

3. krok. Riešenie hľadáme v tvare trojprvkovej množiny $\Delta = \{i < j < k\}$, kde $i, j, k \in \{1, 2, ..., p\}$. Podobne ako v predchádzajúcom kroku toto riešenie je zahrnuté do množiny riešení $\mathcal{D}$ práve vtedy, ak táto množina neobsahuje také riešenie Δ', ktoré by bolo podmnožinou aktuálneho riešenia $\Delta = \{i < j < k\}$.

....

Ilustračný príklad

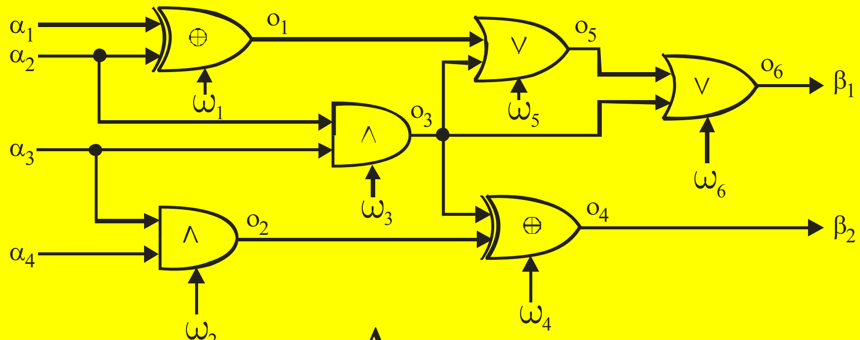
Použitý systém je interpretovaný pomocou Boolovej funkcie

$$f_S(\omega): \{0,1\}^4 \rightarrow \{0,1\}^2$$

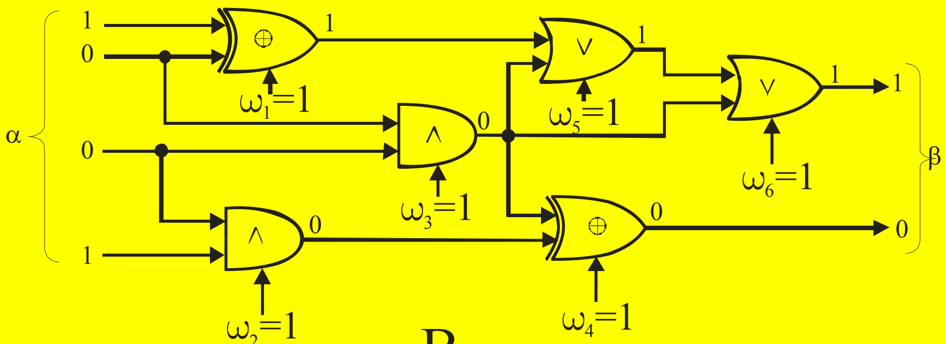
Matematický model tohto systému môžeme jednoducho vytvoriť pomocou diagramu A

$$o_1 = XOR(\alpha_1, \alpha_2; \omega_1),\ o_2 = AND(\alpha_3, \alpha_4; \omega_2),\ o_3 = AND(\alpha_2, \alpha_3; \omega_3)$$
$$\beta_2 = XOR(o_2, o_3; \omega_4),\ o_5 = OR(o_1, o_3; \omega_5),\ \beta_1 = OR(o_3, o_5; \omega_6)$$

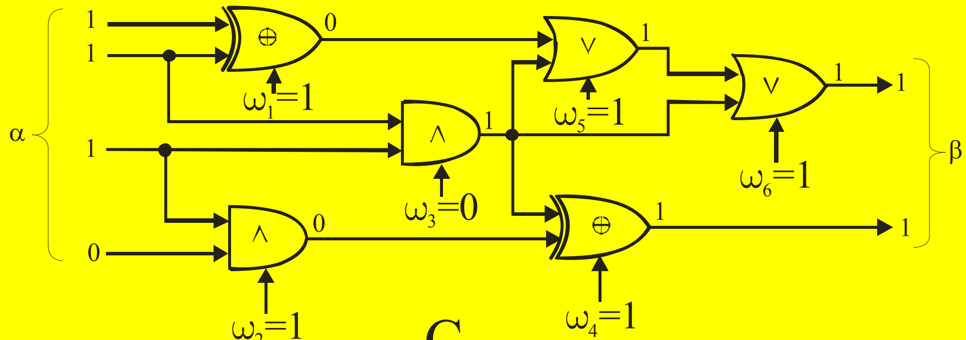
Pomocou týchto jednoduchých formúl môžeme jednoducho spočítať odozvu β na vstup α, pričom pomocou parametrov ω_i môžeme jednoducho modelovať korektnosť alebo nekorektnosť jednotlivých elementárnych logických brán.



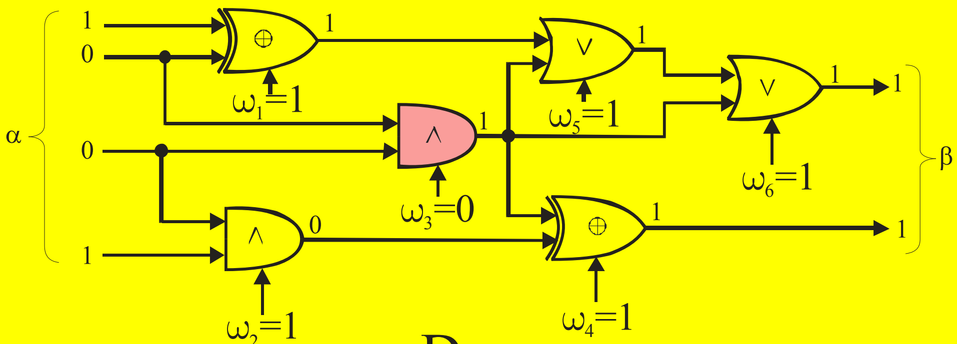
A



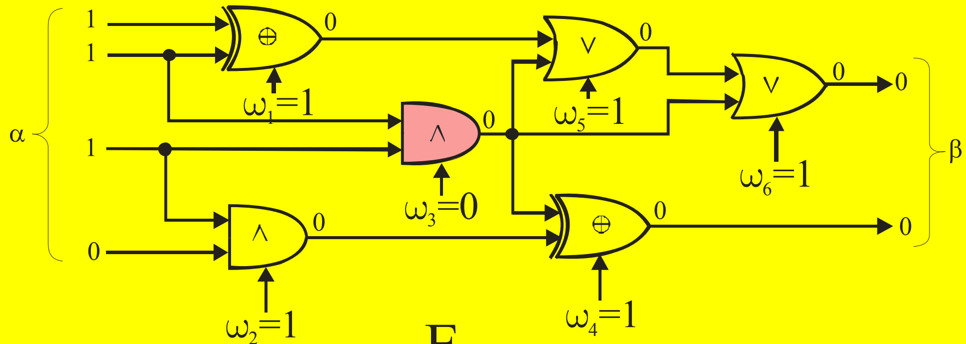
B



C



D



E

Priesvitka 23

Tréningová množina nech má tvar

$$\mathcal{A}_{train} = \left\{ (1001)/(11), (1110)/(00) \right\}$$

Vytvoríme tri účelové funkcie

$$E^{(1)}(\omega) = \left| (11) - f_{\mathcal{S}}\left((1001); \omega \right) \right|,\ E^{(2)}(\omega) = \left| (00) - f_{\mathcal{S}}\left((1110); \omega \right) \right|$$

$$E^{(3)}(\omega) = E^{(1)}(\omega) + E^{(2)}(\omega)$$

Optimálne hodnoty Δ-množiny pre rôzne tvary účelovej funkcie

$E^{(1)}(\omega)$	$E^{(2)}(\omega)$	$E^{(1)}(\omega)+E^{(2)}(\omega)$
$\Delta_2 = \{3\}$	$\Delta_1 = \{3\}$	$\Delta_1 = \{3\}$
$\Delta_1 = \{2\}$	$\Delta_2 = \{2,6\}$	$\Delta_2 = \{1,2,6\}$
$\Delta_3 = \{4\}$	$\Delta_3 = \{4,6\}$	$\Delta_3 = \{1,4,6\}$
		$\Delta_4 = \{2,5,6\}$
		$\Delta_5 = \{4,5,6\}$

Záver

- Použitý prístup k štúdiu určenia diagnózy je aplikovaný na parametrické binárne systémy so vstupom a výstupom, pričom binárne parametre špecifikuje korektnosť-chybovosť elementárnych Boolových funkcií.
- Stanovenie diagnózy sa redukuje na minimalizáciu účelovej funkcie vzhľadom k binárnemu parametru, ktorý špecifikuje korektnosť-chybovost použitých elementárnzch Boolových funkcii.
- K riešeniu optimalizačného problému účelovej funkcie existujú efektívne heuristiky, ktoré podstatne zjednodušujú hľadanie optimálnzch riešení s nulovou hodnotou účelovej funkcie.

Čo ďalej?

- Existuje možnosť maskovania chyby?
- Výsledky interpretovať v rámci kognitívnej vedy, ako kognitívny model pre konštrukciu diagnózy.
- ....

Pán doktor Mäsiar práve prednáša na fakulte. Po pípnutí, povedzte svoje meno, telefónne číslo a krátku diagnózu, keď sa vráti, tak sa na to podíva. "

The End