



CONFEDERAZIONE SVIZZERA
ISTITUTO FEDERALE DELLA PROPRIETÀ INTELLETTUALE

(11) **CH** **711 716 A1**

(51) Int. Cl.: **G06N** 7/00 (2006.01)
G06F 17/00 (2006.01)
G06N 5/00 (2006.01)

Domanda di brevetto per la Svizzera ed il Liechtenstein

Trattato sui brevetti, del 22 dicembre 1978, fra la Svizzera ed il Liechtenstein

(12) **DOMANDA DI BREVETTO**

(21) Numero della domanda:	01583/15	(71) Richiedente:	SUPSI, Via Cantonale 2c Galleria 2 6928 - Manno (CH)
(22) Data di deposito:	29.10.2015	(72) Inventore/Inventori:	Mauro Scanagatta, 6934 Bioggio (CH) Cassio Polpo de Campos, BT9 6SA Belfast (GB) Giorgio Corani, 22100 Como (IT) Marco Zaffalon, 6900 Lugano (CH)
(43) Domanda pubblicata:	15.05.2017	(74) Mandatario:	MODIANO & PARTNERS SA, Via Calloni, 1 6900 Lugano (CH)

(54) **Apprendimento della struttura di reti bayesiane da un insieme di dati completo**

(57) Il presente trovato si riferisce ad un dispositivo (1) per l'apprendimento della struttura di una rete bayesiana relativa ad una pluralità di variabili, ciascuna di dette variabili potendo assumere una pluralità di stati finiti, comprendente:

- un modulo di inizializzazione (30) per il calcolo di punteggi esatti nelle relazioni di parentela tra ciascuna di dette variabili e le rimanenti variabili;
- un motore di calcolo (32) per il calcolo di punteggi euristici nelle relazioni di parentela di ciascuna di dette variabili rispetto alle coppie di dette rimanenti variabili, per la selezione dei sottoinsiemi di genitori aventi punteggio più elevato, e per il calcolo di punteggi euristici relativi a detti sottoinsiemi di genitori integrati con il punteggio relativo ad una di dette rimanenti variabili, generando nuovi sottoinsiemi;
- un modulo di ordinamento (34) configurato per ordinare detti sottoinsiemi di genitori in una lista, in ordine di punteggio;
- un modulo di ottimizzazione (40) configurato per generare iterativamente, per un periodo di tempo selezionabile, una sequenza casuale degli elementi presenti in detta lista di detti sottoinsiemi di genitori e per calcolare un punteggio complessivo relativo all'insieme di genitori che copra l'intero insieme di dette variabili, tenendo traccia dell'insieme di genitori cui corrisponde il punteggio più elevato.

Descrizione

[0001] Il presente trovato ha come oggetto un dispositivo e un metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo particolarmente, seppur non esclusivamente, utili e pratici per l'apprendimento della struttura di reti bayesiane con elevato numero di variabili, utilizzabili per l'analisi e la correlazione di elevate quantità di dati. Il dispositivo ed il metodo di cui al presente trovato si applicano, ad esempio, alla diagnosi di reti mobili, all'analisi di dati finanziari, all'analisi decisionale, alle previsioni meteorologiche e alla predizione di comportamenti in genere.

[0002] Una rete bayesiana è la rappresentazione grafica di un modello probabilistico costituito da un insieme di variabili aleatorie e dalle loro dipendenze condizionali, rappresentate con l'ausilio di un grafo aciclico orientato, definito Directed Acyclic Graph (DAG), detto anche struttura della rete.

[0003] Un grafo aciclico orientato è un grafo dotato di nodi ed archi orientati ma privo di cicli orientati.

[0004] Questo significa che, a partire da qualunque nodo del grafo, non è possibile incrociare nuovamente il nodo di partenza muovendosi secondo il verso degli archi della rete.

[0005] I nodi rappresentano variabili aleatorie, ove ciascuna variabile può assumere un determinato stato o valore.

[0006] Ai possibili stati di ciascuna variabile, che devono essere mutualmente esclusivi, è associato un valore di probabilità, mentre gli archi tra nodi indicano una relazione di dipendenza condizionale tra le variabili rappresentate dai nodi.

[0007] Ai nodi che presentano genitori, cioè che sono collegati ad almeno un arco che punta su di loro, sono associate tabelle di probabilità condizionata, definite Conditional Probability Table (CPT), ossia le tabelle che contengono le probabilità dei valori del nodo condizionate alle possibili combinazioni di valori dei nodi genitori.

[0008] Più in dettaglio, la rappresentazione grafica del modello probabilistico di cui una rete bayesiana permette di rappresentare e analizzare un fenomeno oggetto di studio in condizioni di incertezza. Il modello probabilistico rappresenta una distribuzione di probabilità su un insieme di variabili aleatorie discrete $X = \{X_1/X_2, \dots, X_n\}$.

[0009] Le reti bayesiane, tipicamente denotate con $B = (G, \theta)$, sono definite attraverso la specificazione di due componenti: una componente qualitativa e una componente quantitativa.

[0010] La componente qualitativa consiste appunto nel grafo aciclico orientato, il summenzionato Directed Acyclic Graph (DAG), denotato con $G = (V, \epsilon)$, in cui i nodi V sono in corrispondenza biunivoca con l'insieme $X = \{X_1/X_2, \dots, X_n\}$ di variabili aleatorie discrete e gli archi orientati ϵ sono coppie ordinate di elementi di V . Ogni arco rappresenta appunto la dipendenza condizionata fra i nodi che connette.

[0011] La componente quantitativa consiste, invece, in un insieme di distribuzioni di probabilità condizionate, o Conditional Probability Distribution (CPD), su X, θ , definite parametri della rete. Le distribuzioni di probabilità condizionate locali, associate ad ogni variabile aleatoria e condizionate da ogni possibile combinazione dei valori assunti dall'insieme di genitori della variabile, sono specificate attraverso un insieme di parametri.

[0012] In pratica, la struttura di una rete bayesiana codifica le relazioni qualitative, rappresentate per mezzo di archi, che sussistono tra le variabili aleatorie discrete, rappresentate per mezzo di nodi.

[0013] La forza delle relazioni che esistono tra le variabili aleatorie discrete, ovvero il peso degli archi che mettono in relazione i nodi, è quantificata dalle distribuzioni di probabilità condizionata associate ad ogni nodo.

[0014] Se un nodo ha molti genitori o se i genitori di un nodo hanno numerosi stati possibili, la tabella di probabilità condizionata, o Conditional Probability Table (CPT), associata può essere molto grande. La dimensione di una tabella di probabilità condizionata è infatti esponenziale rispetto al numero dei genitori.

[0015] Il processo di apprendimento o learning di una rete bayesiana comprende sostanzialmente due fasi distinte. La prima fase riguarda l'apprendimento della struttura della rete, o structural learning, ovvero le relazioni fra le variabili. La seconda fase riguarda l'apprendimento dei parametri della rete, o learning parameters, ovvero le probabilità condizionate.

[0016] In particolare, l'apprendimento della struttura di una rete bayesiana a partire da un insieme di dati completo è un problema NP-difficile.

[0017] Attualmente sono noti svariati algoritmi esatti per l'apprendimento della struttura di una rete bayesiana in funzione di un punteggio o score, vale a dire algoritmi aventi il compito di trovare la struttura della rete bayesiana che massimizzi il punteggio o score dipendente dall'insieme di dati. Tali algoritmi sono basati su tecniche differenti note nello stato dell'arte, tra cui ad esempio la programmazione dinamica, la tecnica branch and bound, la programmazione lineare e intera, o l'euristica del percorso più breve.

[0018] Tipicamente l'apprendimento della struttura della rete viene eseguito nel corso di due fasi distinte. Dapprima si procede con una identificazione dell'insieme dei genitori e, successivamente, l'ottimizzazione della struttura.

[0019] La fase di identificazione dell'insieme di genitori produce, per ogni variabile aleatoria, ovvero per ogni nodo, una lista di insiemi di genitori candidati, vale a dire una lista di insiemi di genitori idonei a massimizzare il punteggio della rete bayesiana.

[0020] La fase di ottimizzazione della struttura, invece, seleziona, per ogni nodo, un insieme di genitori tra quelli elencati nella suddetta lista e lo assegna al corrispondente nodo, massimizzando il punteggio della struttura risultante senza introdurre cicli.

[0021] Il tipico problema che si deve affrontare nella fase di identificazione dell'insieme di genitori consiste nel fatto che è improbabile che tale fase ammetta un algoritmo tempo-polinomiale con una garanzia di buona qualità.

[0022] Questo problema ha spinto la ricerca verso lo sviluppo di euristiche di ricerca efficaci. Tuttavia, di norma, il grado di ingresso k , o in-degree, indicante il numero di genitori per nodo, viene definito preventivamente, e poi viene calcolato il punteggio o score di tutti gli insiemi di genitori.

[0023] Si noti che un grado di ingresso maggiore implica uno spazio di ricerca più ampio e permette conseguentemente di ottenere un punteggio maggiore.

[0024] Tali soluzioni note non sono scevre da inconvenienti, tra i quali, in particolare, va annoverato il fatto che esse richiedono un tempo di calcolo elevatissimo, esponenziale rispetto al grado di ingresso definito preventivamente. In pratica, maggiore è il grado di ingresso k , maggiore sarà il tempo di calcolo e il carico computazionale.

[0025] Scegliendo a priori il grado di ingresso, l'utente accetta un compromesso tra due diversi obiettivi: minimizzare il tempo di calcolo e massimizzare il punteggio della rete bayesiana. Tuttavia l'utente può solamente stimare in maniera approssimativa l'impatto del grado di ingresso sul tempo di calcolo e sul punteggio della rete bayesiana. Quando il numero di variabili aleatorie, e quindi di nodi, è molto elevato, il grado di ingresso viene in genere impostato ad un valore piccolo, al fine di mantenere la fattibilità della fase di ottimizzazione della struttura.

[0026] Compito precipuo del presente trovato è quello di superare i limiti dell'arte nota sopra esposti, escogitando un dispositivo e un metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo che consentano di ottenere effetti analoghi o migliori a quelli ottenibili con le soluzioni di tipo noto, permettendo di individuare una struttura di una rete bayesiana di buona qualità, ossia per quanto possibile ottimizzata e con un punteggio massimizzato, mantenendo i tempi di calcolo ad un livello contenuto.

[0027] Nell'ambito di questo compito, uno scopo del presente trovato è quello di concepire un dispositivo e un metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo che permettano di definire preventivamente il tempo di calcolo a disposizione per l'individuazione di una struttura di una rete bayesiana di buona qualità.

[0028] Un altro scopo del presente trovato è quello di escogitare un dispositivo e un metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo che siano facilmente applicabili a insiemi massivi di dati anche con elevato numero di variabili, ad esempio dati finanziari, dati merceologici, dati meteorologici, dati relativi ad accesso a pagine web, e così via.

[0029] Non ultimo scopo del presente trovato è quello di realizzare un dispositivo e un metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo che siano di elevata affidabilità, di relativamente semplice realizzazione ed a costi contenuti.

[0030] Questo compito, nonché questi ed altri scopi che meglio appariranno in seguito, sono raggiunti da un dispositivo per l'apprendimento della struttura di una rete bayesiana relativa ad una pluralità di variabili, ciascuna di dette variabili potendo assumere una pluralità di stati finiti, comprendente:

- un modulo di inizializzazione per il calcolo di punteggi esatti nelle relazioni di parentela tra ciascuna di dette variabili e le rimanenti variabili;
- un motore di calcolo per il calcolo di punteggi euristici nelle relazioni di parentela di ciascuna di dette variabili rispetto alle coppie di dette rimanenti variabili, per la selezione dei sottoinsiemi di genitori aventi punteggio più elevato, e per il calcolo di punteggi euristici relativi a detti sottoinsiemi di genitori integrati con il punteggio relativo ad una di dette rimanenti variabili, generando nuovi sottoinsiemi;
- un modulo di ordinamento configurato per ordinare detti sottoinsiemi di genitori in una lista, in ordine di punteggio;
- un modulo di ottimizzazione configurato per generare iterativamente, per un periodo di tempo selezionabile, una sequenza casuale degli elementi presenti in detta lista di detti sottoinsiemi di genitori e per calcolare un punteggio complessivo relativo all'insieme di genitori che copra l'intero insieme di dette variabili, tenendo traccia dell'insieme di genitori cui corrisponde il punteggio più elevato.

[0031] Il compito e gli scopi prefissati sono altresì raggiunti da un metodo per l'apprendimento della struttura di una rete bayesiana relativa ad una pluralità di variabili, ciascuna di dette variabili potendo assumere una pluralità di stati finiti, comprendente i passi che consistono nel:

- inizializzare i punteggi esatti nelle relazioni di parentela tra ciascuna di dette variabili e le rimanenti variabili;
- calcolare i punteggi euristici nelle relazioni di parentela di ciascuna di dette variabili rispetto alle coppie di dette rimanenti variabili, per la selezione dei sottoinsiemi di genitori aventi punteggio più elevato;
- calcolare i punteggi euristici relativi a detti sottoinsiemi di genitori integrati con il punteggio relativo ad una di dette rimanenti variabili, generando nuovi sottoinsiemi;
- ordinare detti sottoinsiemi di genitori in una lista, in ordine di punteggio;

- ottimizzare la struttura della rete bayesiana generando iterativamente, per un periodo di tempo selezionabile, una sequenza casuale degli elementi presenti in detta lista di detti sottoinsiemi di genitori e calcolando un punteggio complessivo relativo all'insieme di genitori che copra l'intero insieme di dette variabili, tenendo traccia dell'insieme di genitori cui corrisponde il punteggio più elevato.

[0032] Ulteriori caratteristiche e vantaggi del trovato risulteranno maggiormente dalla descrizione di una forma di realizzazione preferita, ma non esclusiva, del dispositivo e del metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo secondo il trovato, illustrata a titolo indicativo e non limitativo negli uniti disegni, in cui:

- la fig. 1 è uno schema a blocchi che illustra schematicamente i componenti di una forma di realizzazione del dispositivo per l'apprendimento della struttura di reti bayesiane, secondo il presente trovato;
- la fig. 2 rappresenta una nota rete bayesiana costruita manualmente da esperti in biologia in relazione alla diagnosi di malattie polmonari nei bambini;
- la fig. 3 rappresenta una rete bayesiana generata dal dispositivo e dal metodo secondo il presente trovato a partire da un insieme di dati completo, confrontabile con la rete di fig. 2.

[0033] Con riferimento alle figure citate, il dispositivo, indicato globalmente con il numero di riferimento 1, e il metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo secondo il trovato verranno ora illustrati in relazione ai passi che compongono tale metodo.

[0034] Di seguito viene descritto il primo passo del metodo per l'apprendimento della struttura di reti bayesiane secondo il presente trovato. Il passo dell'identificazione dell'insieme di genitori (Parent Set Identification) secondo il presente trovato, che utilizza la valutazione euristica del punteggio o score della struttura di una rete bayesiana, è attuato attraverso un modulo di identificazione 20.

[0035] Nell'ambito di questo primo passo sono impiegate principalmente due liste 10, 12.

[0036] La prima lista 10 è definita lista open, ed è una lista in cui sono elencati gli insiemi di genitori ancora da esplorare, ordinati sulla base del loro punteggio euristico.

[0037] La seconda lista 12 è definita lista closed, e rappresenta una lista in cui sono elencati gli insiemi già esplorati, insieme al loro punteggio esatto.

[0038] Sono poi utilizzate le seguenti funzioni: una prima funzione 21 score(P, X), che calcola un punteggio esatto secondo metodologie note tra loro alternative, ad esempio un punteggio BDeu (Bayesian Dirichlet equivalent uniform) oppure un punteggio BIC (Bayesian information criterion) esatto per l'insieme di variabili P come genitori della variabile X; una seconda funzione 22 c(Y, sk) che memorizza in un'area di memoria temporanea, o cache, il punteggio per una variabile Y come singolo genitore;

una terza funzione 23 s(Y), che recupera il punteggio per la variabile X come singolo genitore;

una quarta funzione 24 pop(L), che estrae l'insieme di genitori con il punteggio migliore dalla lista L; una quinta funzione 25 add(L, P, s), che aggiunge l'insieme di genitori P alla lista L, con il punteggio s;

una sesta funzione 26 time(), che restituisce il valore booleano vero o true se c'è ancora tempo disponibile per il calcolo computazionale.

[0039] Essendo a conoscenza di una variabile obiettivo X e di un insieme di variabili candidate Y, lo scopo del modulo di identificazione 20 dell'insieme di genitori è quello di trovare il sottoinsieme, o subset, di Y che produce i punteggi o score migliori come insieme di genitori per X.

[0040] Come primo passo, un modulo di inizializzazione 30 calcola il punteggio esatto per ogni variabile candidata Y come singolo genitore di X con punteggio $\{Y, X\}$, aggiunge l'insieme e il punteggio alla lista closed 12, e salva in memoria cache il risultato per consentire un recupero dello stesso in un momento futuro.

[0041] Successivamente, il modulo di inizializzazione 30 procede con l'aggiunta alla lista open 10 di tutte le coppie di variabili candidate Y_1 e Y_2 , assegnando ad esse il punteggio euristico ottenuto come $s(Y_1, X) + s(Y_2, X)$.

[0042] A questo punto, la lista open 10 risulta inizializzata e un motore di calcolo 32 può procedere con l'esecuzione del ciclo principale, ripetendo i seguenti passi finché tutti gli elementi nella lista open 10 sono stati processati, oppure finché il tempo disponibile per il calcolo computazionale è finito.

[0043] Inizialmente, il motore di calcolo 32 estrae dalla lista open 12 il sottoinsieme P con punteggio euristico migliore, calcola il punteggio esatto per l'insieme di genitori relativo a detto sottoinsieme P, aggiungendo l'insieme e il punteggio alla lista closed 10.

[0044] Il motore di calcolo 32 procede poi ricercando tutte le possibili espansioni di P che risultano dall'aggiunta di una singola variabile candidata Y, previa verifica che la candidata Y non facesse già parte di un sottoinsieme P già trattato, ossia che l'insieme di variabili $P \cup Y$ non esista già all'interno della lista open 12 o della lista closed 10.

CH 711 716 A1

[0045] Il motore di calcolo 32 assegna allora un punteggio euristico ottenuto come $s(P,X) + s(Y,X)$ a ciascuno di questi insiemi di genitori candidati, e aggiunge questi insiemi alla lista open 12.

[0046] Infine, un modulo di ordinamento 34 dell'insieme di genitori restituisce il contenuto della lista closed 10, ordinato in base al punteggio esatto precedentemente calcolato.

[0047] Le operazioni effettuate dal modulo di identificazione 20 sono per maggior chiarezza sintetizzate nelle seguenti righe di pseudo-codice.

PARENT SET IDENTIFICATION

```
1:  function search(children variable x)
2:  open  $\leftarrow \emptyset$ 
3:  closed  $\leftarrow 0$ 
4:  for all candidate variable Y do
5:      sk  $\leftarrow$  score ( {Y}, X)
6:      C(Y, sk)
7:      add(closed, Y, sk)
8:  end for
9:  for all pair of candidate variables  $Y_1$  and  $Y_2$  do
10:     h  $\leftarrow$  s( $Y_1,X$ ) + s( $Y_2,X$ )
11:     add(open, { $Y_1, Y_2$ }, h)
12  end for
13:  while open  $\neq \emptyset$  & time ( ) do
14:     P  $\leftarrow$  pop (open)
15:     sk  $\leftarrow$  score(P, X)
16:     for all candidate variable  $Y \in P, P \cup Y \in$ 
       open  $\cup$  closed do
17:         h  $\leftarrow$  sk + s(Y,X)
18:         add(open, P  $\cup$  Y, h)
19:     end for
20:  end while
       return ordered(closed)
21:  end function
```

[0048] Di seguito viene descritto il secondo passo del metodo per l'apprendimento della struttura di reti bayesiane secondo il presente trovato. Il passo dell'ottimizzazione della struttura di una rete bayesiana (Structure Optimization) secondo il presente trovato è attuato attraverso un modulo di ottimizzazione 40.

[0049] Il modulo di ottimizzazione 40 implementa le seguenti funzioni:

una prima funzione 42 ancestor(o, p, a), che dato un vettore o array a degli ascendenti, controlla se p è ascendente di o;

una seconda funzione 44 update(o, p, a), che aggiorna l'array a degli ascendenti o discendenti con l'insieme di genitori p per o;

una terza funzione 46 score (p), che calcola il punteggio dell'insieme di genitori p;

una quarta funzione 48 randomize(L), che pone in ordine casuale gli elementi della lista L;

una quinta funzione 50 time(), che restituisce vero o true se c'è ancora tempo disponibile per il calcolo computazionale.

[0050] Essendo a conoscenza, per ogni variabile o nodo, di una lista degli insiemi di genitori più promettenti, assieme ai relativi punteggi, lo scopo dal modulo di ottimizzazione 40 della struttura è quello di trovare ed assegnare il miglior insieme di genitori possibile ad ogni variabile o nodo.

[0051] Si rammenta che, secondo il trovato, è necessario assicurare l'aciclicità del grafo, ovvero nessun ciclo orientato può essere presente nel grafo, nonché massimizzare il punteggio o score risultante, il quale può essere decomposto come la somma dei punteggi dei singoli insiemi di genitori.

[0052] Al fine di comprendere appieno il controllo aciclico, può essere utile pensare in termini di ascendenti di una variabile, cioè i suoi genitori, i suoi nonni, e così via, fino alle radici del grafo. Infatti, si introduce un ciclo solo quando una variabile o nodo viene connessa a uno dei suoi ascendenti nel grafo.

[0053] Per assicurare il rispetto di tale vincolo, vale a dire per assicurare l'assenza di cicli orientati nel grafo, il modulo di ottimizzazione 40 può utilizzare, in una forma di realizzazione particolarmente efficace, un array di bit, in particolare un array di bit per variabile o nodo. Data la variabile i , l'array di bit per la variabile i contiene, per ogni posizione j , l'informazione indicante se la variabile j è un ascendente della variabile i nel grafo o meno.

[0054] Il modulo di ottimizzazione 40 ripete la stessa operazione per i discendenti di una variabile, ricorsivamente per tutti i figli della variabile stessa. Il modulo di ottimizzazione 40 utilizza l'array dei discendenti per aggiornare l'array degli ascendenti in seguito a ciascun assegnamento di un insieme di genitori, questo assegnamento potendo essere effettuato mediante operazioni binarie molto rapide in termine di esecuzione.

[0055] Il ciclo principale del metodo secondo il trovato può applicare il Metodo Monte Carlo, noto allo stato dell'arte, sui possibili ordinamenti delle variabili o nodi. Ad ogni iterazione, il modulo di ottimizzazione 40 esegue una permutazione casuale, selezionando un nuovo ordinamento.

[0056] In seguito alla selezione di questo ordine, il modulo di ottimizzazione 40 seleziona, per ogni variabile, gli insiemi di genitori con punteggio migliore e che allo stesso tempo non introducono cicli. Alla fine dell'iterazione, il modulo di ottimizzazione confronta la struttura risultante, vale a dire la scelta di un insieme di genitori per ogni variabile o nodo, con la struttura migliore individuata fino a quel momento. Quando il tempo a disposizione è terminato, il modulo di ottimizzazione 40 restituisce la struttura con il miglior punteggio che è stata trovata.

[0057] Il motore 40 ritiene ammissibile un insieme di genitori p per una variabile, se nessuna delle variabili genitore è ascendente di x , ossia l'insieme di genitori non introduce cicli orientati nel grafo. Questa verifica può essere effettuata rapidamente ove implementata mediante operazioni binarie su un array di ascendenti.

[0058] Quando il modulo di ottimizzazione 40 ha completato la scelta dell'insieme di genitori, si occupa di aggiornare gli array degli ascendenti e dei discendenti.

[0059] Per questo scopo, il modulo di ottimizzazione 40 imposta la variabile x e i suoi discendenti, così come i discendenti di ogni genitore in p e i suoi ascendenti. Successivamente, imposta ogni p sui suoi ascendenti, così come gli ascendenti di x e i suoi discendenti. Anche in questo caso queste operazioni possono essere effettuate velocemente mediante operazioni binarie.

[0060] Le operazioni effettuate dal modulo di ottimizzazione 40 sono per maggior chiarezza riassunte nelle seguenti righe di pseudo-codice.

STRUCTURE OPTIMIZATION

```

1:  function search(parent sets P)
2:  order ← 1 : N
3:  bestStructure ← nil
4:  bestScore ← -∞
5:  while time() do
6:      randomize(order)
7:      structure ← empty
8:      score ← 0
9:      for all o ∈ order do
10:         p ← bestParent (o, P, ancestors, descendants)
11:         structure(o) ← p

```

STRUCTURE OPTIMIZATION

```

12:         score+ = score(p)
13:     end for
14:     if score > bestScore then
15:         bestScore ← score
16:         bestStructure ← structure
17:     end if
18: end while
    return bestStructure
19: end function

1: function bestParent(variable o, parent sets P,
    ancestors array a, descendants array d)
2: best ← ∅
3: for all parentSet ∈ P(o) do
4:     admissible ← true
5:     for all parent ∈ parentSet do
6:         if ancestor(o, parent, a) then
7:             admissible ← false
8:         end if
9:     end for
10:    if admissible then
11:        best ← parentSet
12:        break
13:    end if
14: end
    for
15: update(ancestor, o, p)
16: update(descendants, o, p)
    return best
17: end function

```

[0061] Si è in pratica constatato come il trovato assolva pienamente il compito e gli scopi prefissati. In particolare, si è visto come il dispositivo e il metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo così concepiti permettono di superare i limiti qualitativi dell'arte nota, in quanto consentono di individuare una struttura di una rete bayesiana di buona qualità, ossia per quanto possibile ottimizzata e con un punteggio massimizzato, mantenendo i tempi di calcolo ad un livello contenuto.

[0062] Un altro vantaggio del dispositivo e del metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo consiste nel fatto che essi permettono di definire preventivamente il tempo di calcolo a disposizione per l'individuazione di una struttura di una rete bayesiana di buona qualità.

[0063] Un ulteriore vantaggio del dispositivo e del metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo consiste nel fatto che essi sono facilmente applicabili a insiemi massivi di dati anche con elevato numero di variabili, ad esempio dati finanziari, dati merceologici, dati meteorologici, dati relativi ad accesso a pagine web, e così via.

[0064] Benché il dispositivo e il metodo per l'apprendimento della struttura di reti bayesiane da un insieme di dati completo secondo il trovato siano stati concepiti in particolare per l'apprendimento della struttura di reti bayesiane con elevato numero di variabili, utilizzabili per l'analisi e la correlazione di elevate quantità di dati, essi potranno comunque essere utilizzati, più generalmente, per l'apprendimento della struttura di reti bayesiane con un qualsiasi numero di variabili, anche meno di una decina di variabili se del caso.

[0065] Il trovato, così concepito, è suscettibile di numerose modifiche e varianti, tutte rientranti nell'ambito del concetto inventivo. Inoltre, tutti i dettagli potranno essere sostituiti da altri elementi tecnicamente equivalenti.

[0066] In pratica, i materiali impiegati, nonché le dimensioni e le forme contingenti, potranno essere qualsiasi a seconda delle esigenze e dello stato della tecnica.

[0067] In conclusione, l'ambito di protezione delle rivendicazioni non deve essere limitato dalle illustrazioni o dalle forme di realizzazione preferite illustrate nella descrizione sotto forma di esempi, ma piuttosto le rivendicazioni devono comprendere tutte le caratteristiche di novità brevettabile che risiedono nella presente invenzione, incluse tutte le caratteristiche che sarebbero trattate come equivalenti dal tecnico del ramo.

Rivendicazioni

1. Dispositivo (1) per l'apprendimento della struttura di una rete bayesiana relativa ad una pluralità di variabili, ciascuna di dette variabili potendo assumere una pluralità di stati finiti, comprendente:
 - un modulo di inizializzazione (30) per il calcolo di punteggi esatti nelle relazioni di parentela tra ciascuna di dette variabili e le rimanenti variabili;
 - un motore di calcolo (32) per il calcolo di punteggi euristici nelle relazioni di parentela di ciascuna di dette variabili rispetto alle coppie di dette rimanenti variabili, per la selezione dei sottoinsiemi di genitori aventi punteggio più elevato, e per il calcolo di punteggi euristici relativi a detti sottoinsiemi di genitori integrati con il punteggio relativo ad una di dette rimanenti variabili, generando nuovi sottoinsiemi;
 - un modulo di ordinamento (34) configurato per ordinare detti sottoinsiemi di genitori in una lista, in ordine di punteggio;
 - un modulo di ottimizzazione (40) configurato per generare iterativamente, per un periodo di tempo selezionabile, una sequenza casuale degli elementi presenti in detta lista di detti sottoinsiemi di genitori e per calcolare un punteggio complessivo relativo all'insieme di genitori che copra l'intero insieme di dette variabili, tenendo traccia dell'insieme di genitori cui corrisponde il punteggio più elevato.
2. Dispositivo (1) per l'apprendimento della struttura di una rete bayesiana secondo la rivendicazione 1, caratterizzato dal fatto che detti punteggi esatti sono calcolati mediante la metodologia BDeu, Bayesian Dirichlet equivalent uniform, oppure la metodologia BIC, Bayesian information criterion, o altre metodologie che funzionano secondo i medesimi principi.
3. Metodo per l'apprendimento della struttura di una rete bayesiana relativa ad una pluralità di variabili, ciascuna di dette variabili potendo assumere una pluralità di stati finiti, comprendente i passi che consistono nel:
 - inizializzare i punteggi esatti nelle relazioni di parentela tra ciascuna di dette variabili e le rimanenti variabili;
 - calcolare i punteggi euristici nelle relazioni di parentela di ciascuna di dette variabili rispetto alle coppie di dette rimanenti variabili, per la selezione dei sottoinsiemi di genitori aventi punteggio più elevato;
 - calcolare i punteggi euristici relativi a detti sottoinsiemi di genitori integrati con il punteggio relativo ad una di dette rimanenti variabili, generando nuovi sottoinsiemi;
 - ordinare detti sottoinsiemi di genitori in una lista, in ordine di punteggio;
 - ottimizzare la struttura della rete bayesiana generando iterativamente, per un periodo di tempo selezionabile, una sequenza casuale degli elementi presenti in detta lista di detti sottoinsiemi di genitori e calcolando un punteggio complessivo relativo all'insieme di genitori che copra l'intero insieme di dette variabili, tenendo traccia dell'insieme di genitori cui corrisponde il punteggio più elevato.
4. Metodo per l'apprendimento della struttura di una rete bayesiana secondo la rivendicazione 3, caratterizzato dal fatto che detti punteggi esatti sono calcolati mediante la metodologia BDeu, Bayesian Dirichlet equivalent uniform, oppure la metodologia BIC, Bayesian information criterion, o altre metodologie che funzionano secondo i medesimi principi.

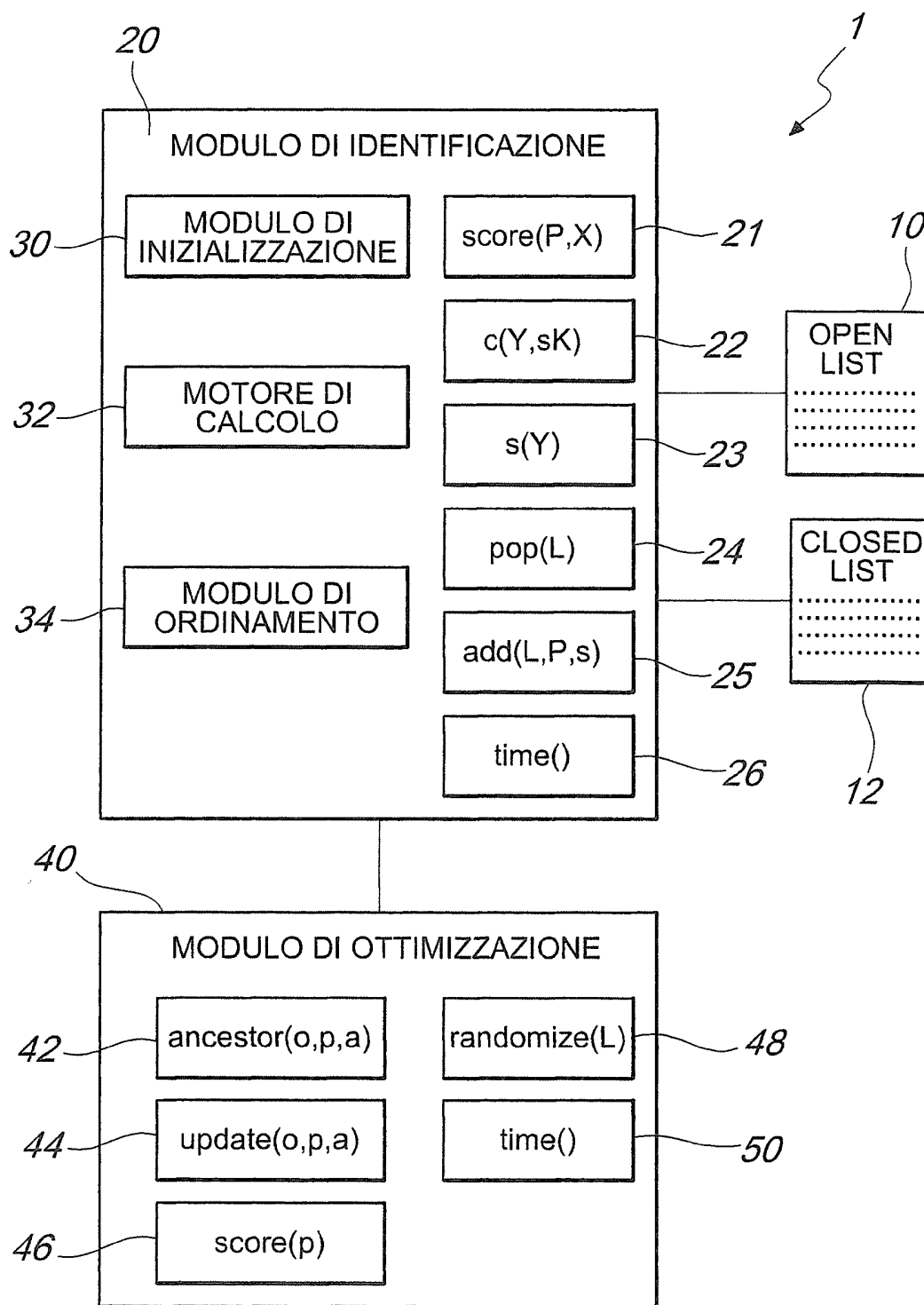


Fig. 1

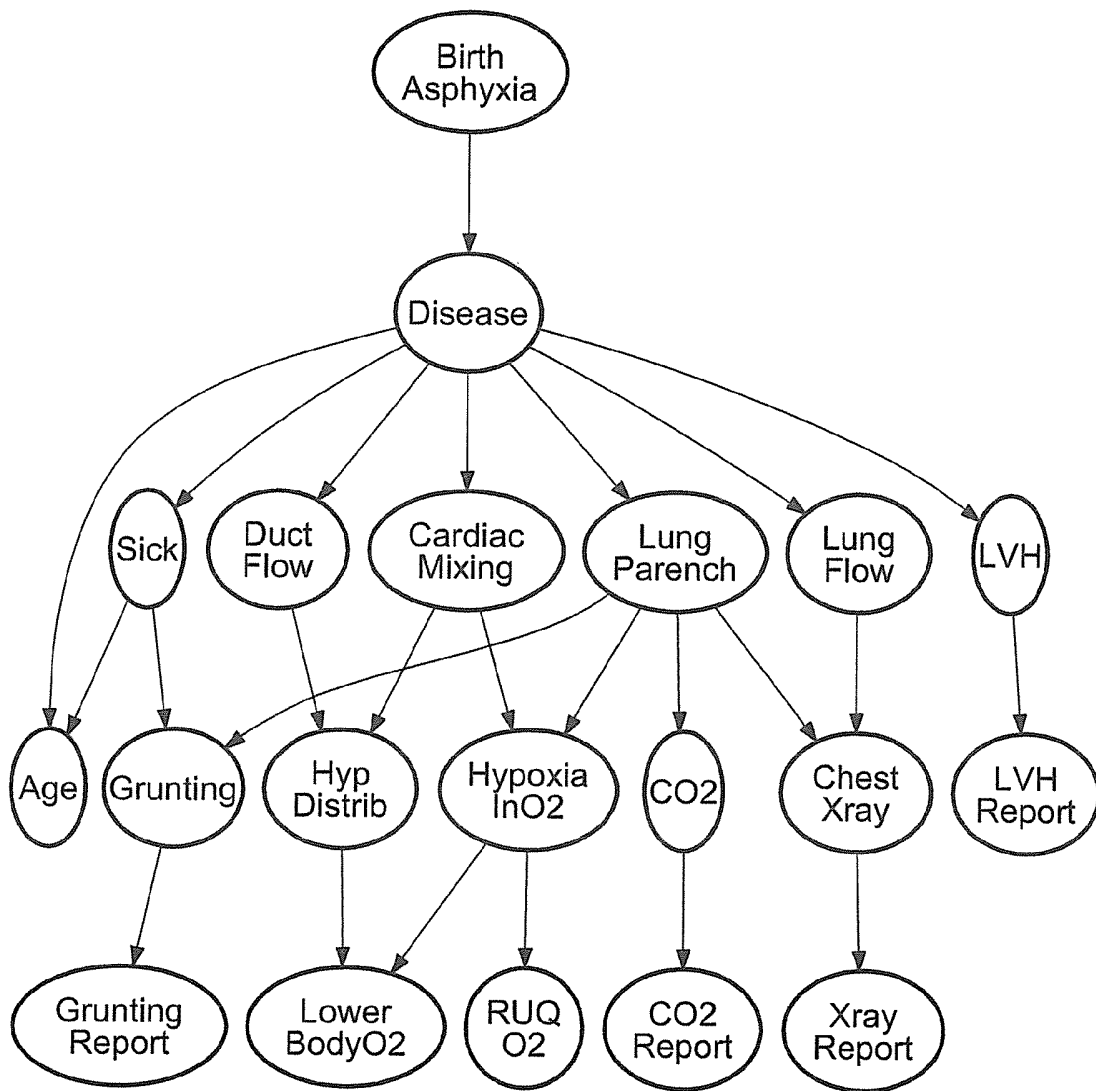


Fig. 2

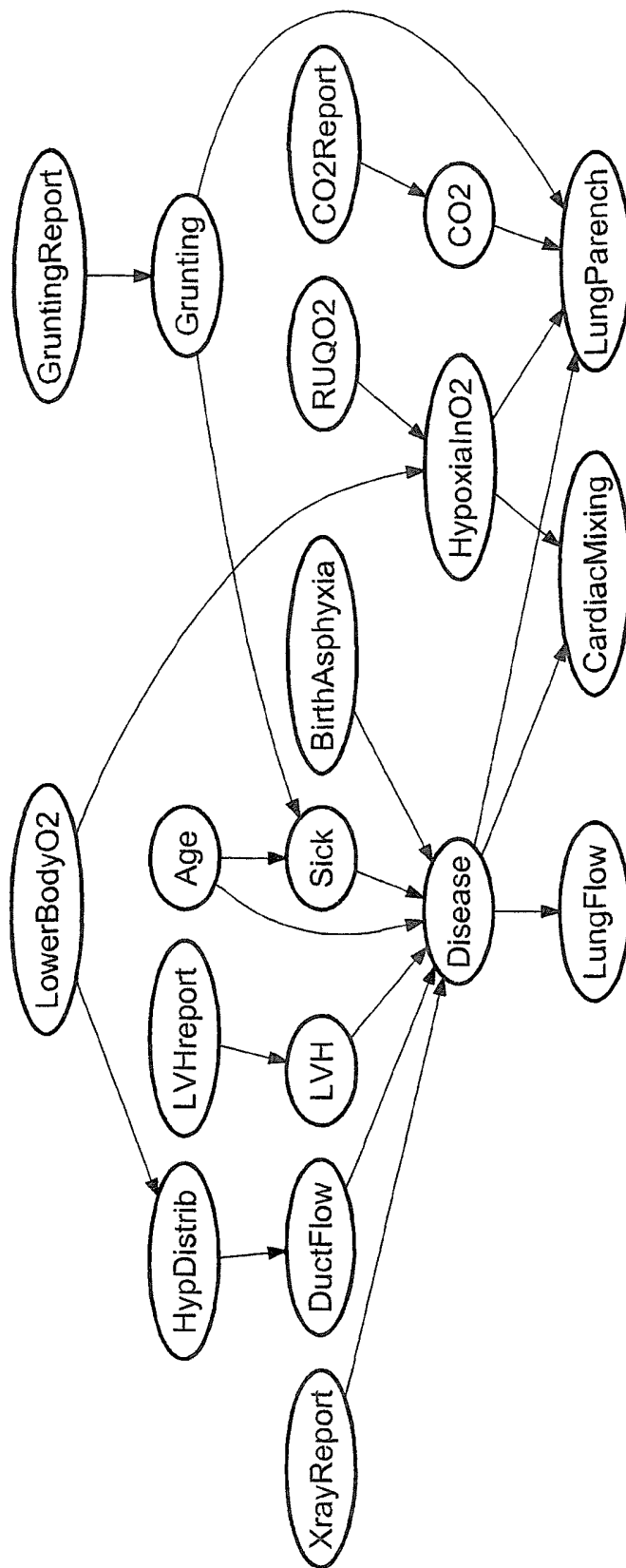


Fig. 3

**RAPPORTO DI RICERCA SULLA
DOMANDA DI BREVETTO SVIZZERA**

Numero della domanda: CH01583/15

Classificazione della domanda (CIB):
G06N7/00, G06F17/00, G06N5/00**Settori ricercati (CIB):**
G06N, G06F**DOCUMENTI RILEVANTI:**

(referenza del documento, categoria, rivendicazioni interessate, indicazione delle parti determinanti (*))

- 1 US2015142709 A1 (SIKORSKY AIRCRAFT CORP [US]) 21.05.2015
 Categoria: **X** Rivendicazioni: **1 - 4**
 * [0004] - [0005]; [0007] - [0009]; [0012]; [0029] - [0031]; [0034] - [0035] *
- 2 US7324981 B2 (MICROSOFT CORP [US]) 27.11.2003
 Categoria: **A** Rivendicazioni: **1 - 4**
 * Colonna 1, linee 7- 9; colonna 2, linee 27 - 39; colonna 5, linee 50 - 56; colonna 22, linee 38 - 41; *
- 3 US2008228676 A1 (FUJI XEROX CO LTD [JP]) 18.09.2008
 Categoria: **A** Rivendicazioni: **1, 3**
 * [0006] *
- 4 US2012185424 A1 (VAIDYANATHAN AKHILESWAR GANESH [US]; JEAN ERIC N [US]; THOMAS MANI [US]; HAMPLE DAVID LOUIS [US]; MCGOWAN MICHAEL THOMAS [US]; WANG JIJUN [US]; FAULKNER ELI T [US]; ASKREN JAY DEE [US]; BOEHLER ALBERT JOSEF [US]; FRAZER DURBAN A [US]; QUANTUM LEAP RES INC [US]) 19.07.2012
 Categoria: **A** Rivendicazioni: **1, 3**
 * [0044] - [0045]; [0060] - [0061]; Figura 4 *
- 5 US2014358831 A1 (HARVARD COLLEGE [US]; GOVERNING COUNCIL OF THE UNIV OF TORONTO THE MARS CT [CA]) 04.12.2014
 Categoria: **A** Rivendicazioni: **1, 3**
 * [0061] - [0065]; [0067]; [0069] *

CATEGORIA DEI DOCUMENTI CITATI:

X:	presi singolarmente mettono in dubbio la novità e/o l'attività inventiva	D:	sono stati citati dal depositante nella domanda
Y:	mettono in dubbio l'attività inventiva in combinazione con un documento della stessa categoria	T:	teorie o principi su cui poggia l'invenzione
A:	definiscono lo stato generale della tecnica senza essere particolarmente rilevanti per la novità e l'attività inventiva	E:	documenti di brevetto la cui data di deposito o di priorità precede la data di deposito della domanda ricercata, ma pubblicati dopo tale data
O:	divulgazione non scritta	L:	documenti citati per altri motivi
P:	sono stati pubblicati tra la data di deposito della domanda di brevetto ricercata e la data di priorità rivendicata	&:	appartenente alla stessa famiglia di brevetti, documento corrispondente

La ricerca si basa sulle rivendicazioni nella versione depositata inizialmente. Una versione modificata delle rivendicazioni depositata in un secondo tempo (art. 51, cpv. 2 OBI) non viene presa in considerazione.

Il presente rapporto di ricerca tiene conto di tutte le rivendicazioni per le quali le dovute tasse sono stata pagate.

Ricercatore: Cristina Mayor
Autorità di ricerca, luogo: Istituto Federale della Proprietà Intellettuale, Berna
Ricerca conclusa il: 22.09.2016

TABELLA DELLE FAMIGLIE DI BREVETTI CITATI

I membri della famiglia di brevetti sono elencati secondo la banca dati dell'Ufficio europeo dei brevetti. L'Ufficio europeo dei brevetti e l'Istituto della Proprietà Intellettuale non garantiscono la correttezza dei dati riportati che servono solo a titolo informativo.

CH 711 716 A1

US2015142709 A1	21.05.2015	US2015142709 A1	21.05.2015
US7324981 B2	27.11.2003	US2003220906 A1	27.11.2003
		US7324981 B2	29.01.2008
US2008228676 A1	18.09.2008	US2008228676 A1	18.09.2008
		JP2008226091 A	25.09.2008
		JP5135831 B2	06.02.2013
US2012185424 A1	19.07.2012	US2012185424 A1	19.07.2012
US2014358831 A1	04.12.2014	KR20160041856 A	18.04.2016
		CA2913743 A1	04.12.2014
		EP3000053 A2	30.03.2016
		JP2016523402 A	08.08.2016
		WO2014194161 A2	04.12.2014
		WO2014194161 A3	29.01.2015
		US2014358831 A1	04.12.2014