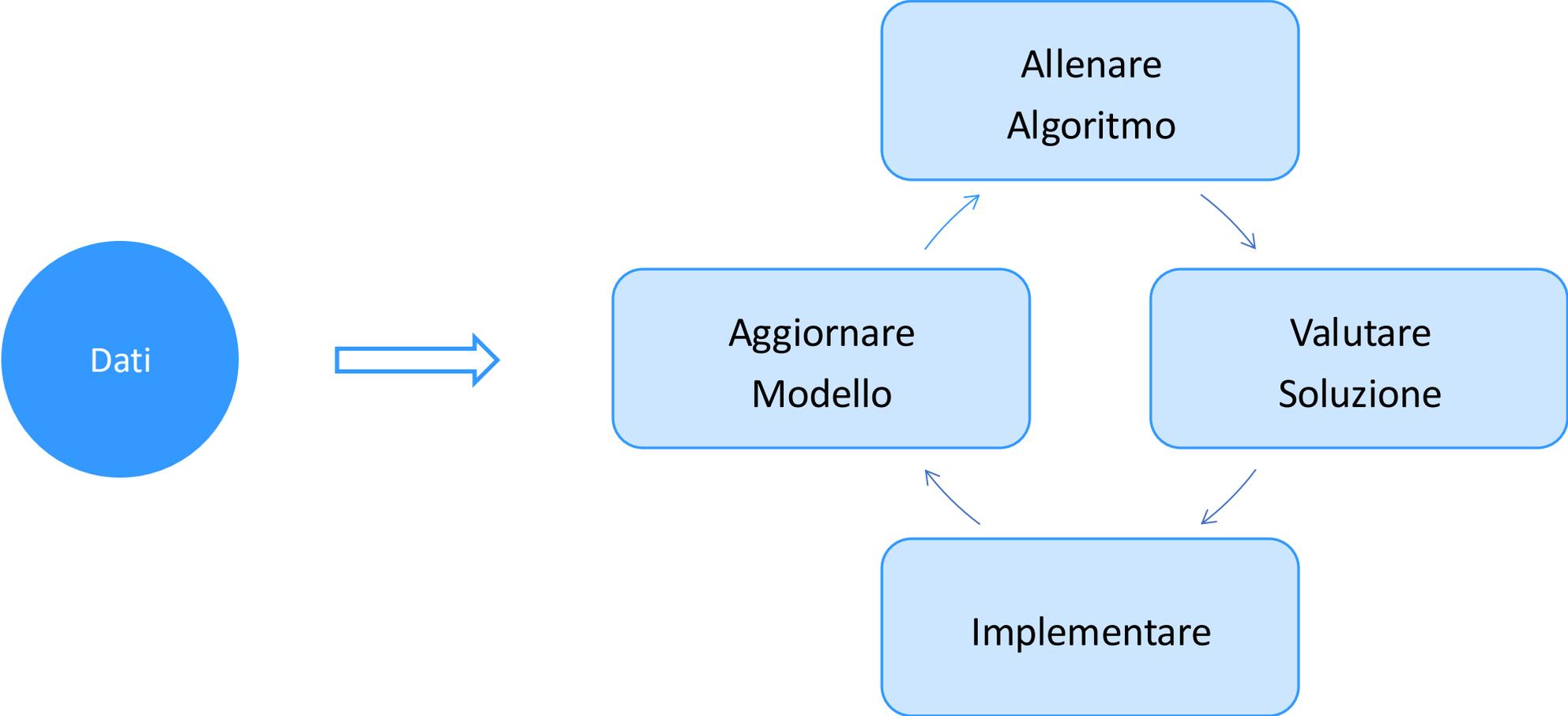


Meti

26/11/2024 - Causal
Inference

POWERED BY  AISENT

Machine Learning: esempio minimale

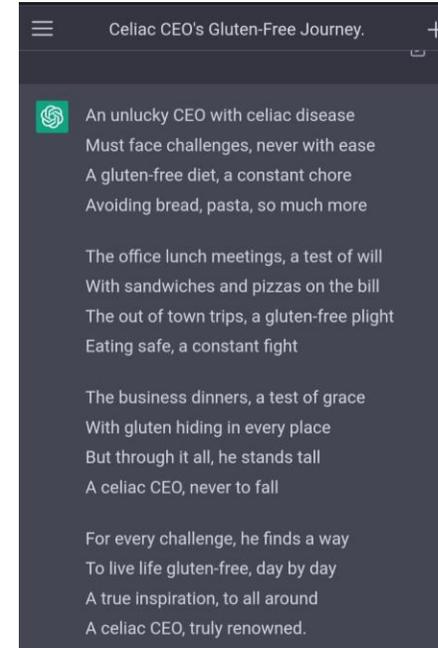


Dove siamo adesso

Gli algoritmi di AI sono sempre più complessi e ottengono risultati sempre migliori e sorprendenti.

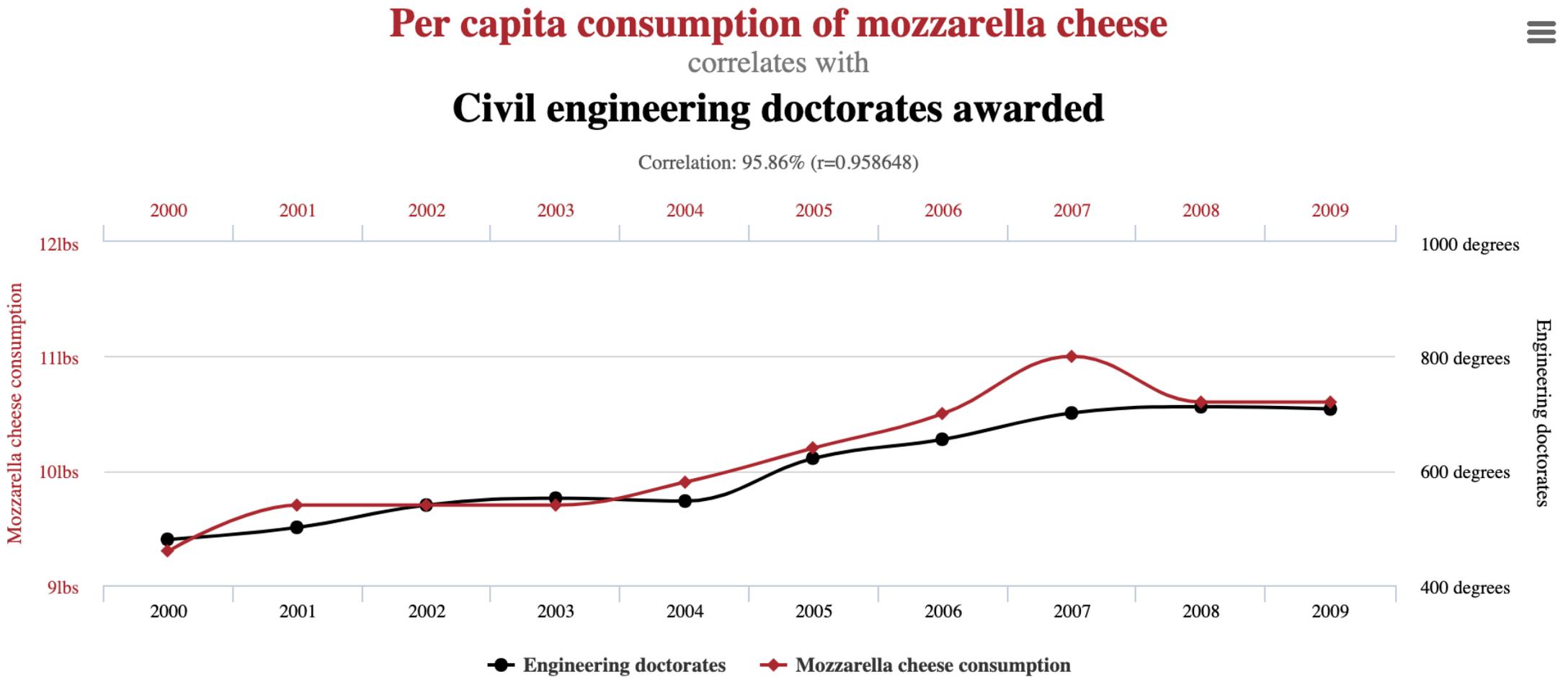
Eppure **nessuno** degli attuali algoritmi può veramente dirsi **intelligente**. Apprendono correlazioni tra variabili, anche molto distanti, ma non riescono a comprendere la relazione tra causa ed effetto.

Come mai non basta la correlazione?



ChatGPT / Midjourney

Correlazione: implica causalità? Non sempre!



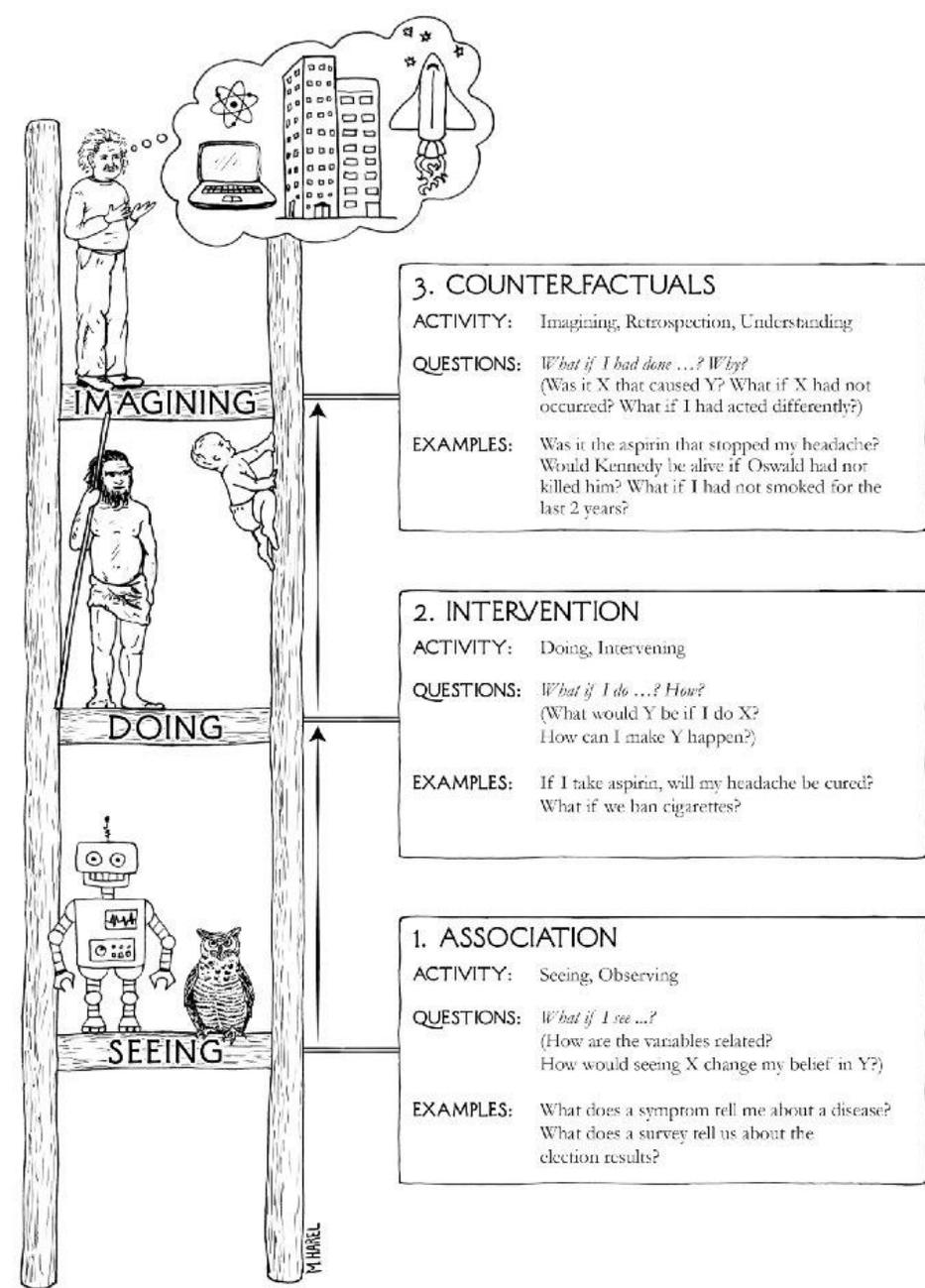
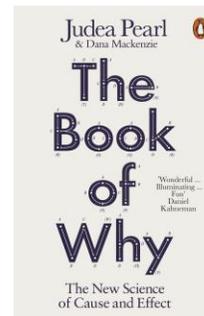
Fonte: <https://www.tylervigen.com/spurious-correlations>

Riprendendo la causalità

I modelli causali partono dalla concezione che gli uomini ragionino su relazioni causa-effetto.

L'unico modo di comprendere veramente quello che succede è sperimentare cosa succede tramite «intervention».

Judea Pearl (UCLA), J. Robins (Harvard), Y. Bengio (Turing Award) ed altri hanno scoperto una serie di tecniche che garantiscono di poter rispondere a quesiti in un contesto causa-effetto.



Di cosa parliamo quando parliamo di “Causal Inference”?

Più nel concreto, possiamo definire il Causal Inference come

- Un insieme di tecniche che ci permettono di capire i rapporti di causa-effetto tra diversi elementi del sistema sotto studio.
- Questo include sia la determinazione della causalità che la sua quantificazione.

Possiamo pensare alla Causal inference come il processo che ci permette di trasformare delle domande o intuizione sul sistema sotto studio in espressioni matematiche in termini di variabili randomiche.

Oggi vedremo diversi modi in cui il Machine Learning ci può aiutare in questo compito.

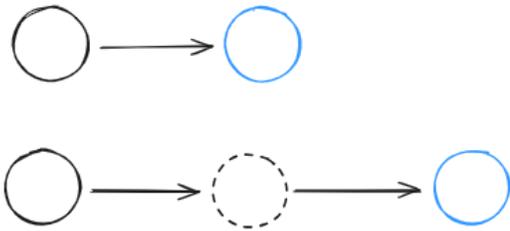
Quando serve?

- Immaginiamo i dati di occupazione di un albergo: il prezzo è positivamente correlato col numero di camere prenotate. Un modello semplice potrebbe concludere che per aumentare l'occupazione basta alzare i prezzi delle stanze.
- Invece noi umani che ragioniamo causalmente sappiamo che nelle stagioni più alte i prezzi delle camere salgono e quindi non è il prezzo a causare una maggiore occupazione.
- Purtroppo gli algoritmi predittivi per costruzione non possono cogliere questi rapporti e bisogna quindi impostarli a mano come nel esempio sopra.
- Ma cosa succede se l'esempio è meno banale degli alberghi nel tempo di vacanze? Se non sappiamo di primi principi i collegamenti causali tra le variabili del nostro sistema rischiamo di basare le nostre predizioni su rapporti causa effetto spuri.

Come funziona?

Esistono tanti tipi di (cor)relazioni tra i dati

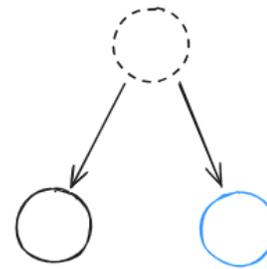
Causation



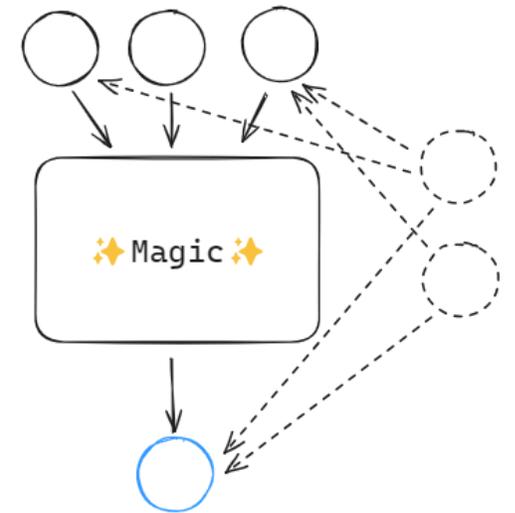
Spurious correlations



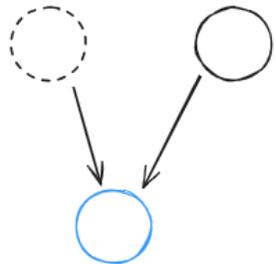
Confounders



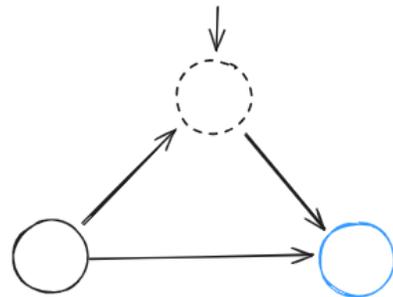
ML Approach



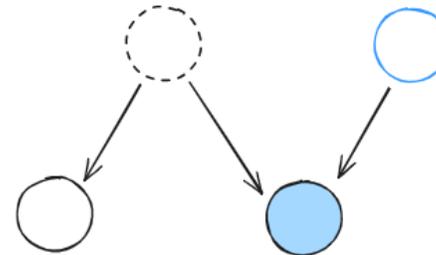
Partial observation



Intermediate effects



Observed variables



Causal Models

Causal Models: Definizione

Come abbiamo detto prima, il concetto di **intervento** è cruciale per dare un senso alla causalità. **Se tutto fosse predeterminato non avrebbe molto senso parlare di causalità!**

Questo nesso si palesa nella frase

“Se io faccio X, allora Y avverrà”

(Da non confondere con *“Se X, allora Y”*).

La definizione più succinta è che un modello Causale ci permette di quantificare l'effetto di una intervento. Ci sono diverse tipologie di modelli, ed in alcune di queste il concetto di intervento sarà più o meno esplicito. Nonostante ciò aiuta a fissare le idee pensare in un modello causale come un strumento per passare da una domanda sul nostro sistema ad una risposta in termini di dati osservazionali o altre informazioni a nostra disposizione.

Causal Models: Intervenzione

$$P(Y | X = x_0)$$

Observational Distribution

Probability of Y given variable X is observed to be value x_0

$$P(Y | do(X = x_0))$$

Interventional Distribution

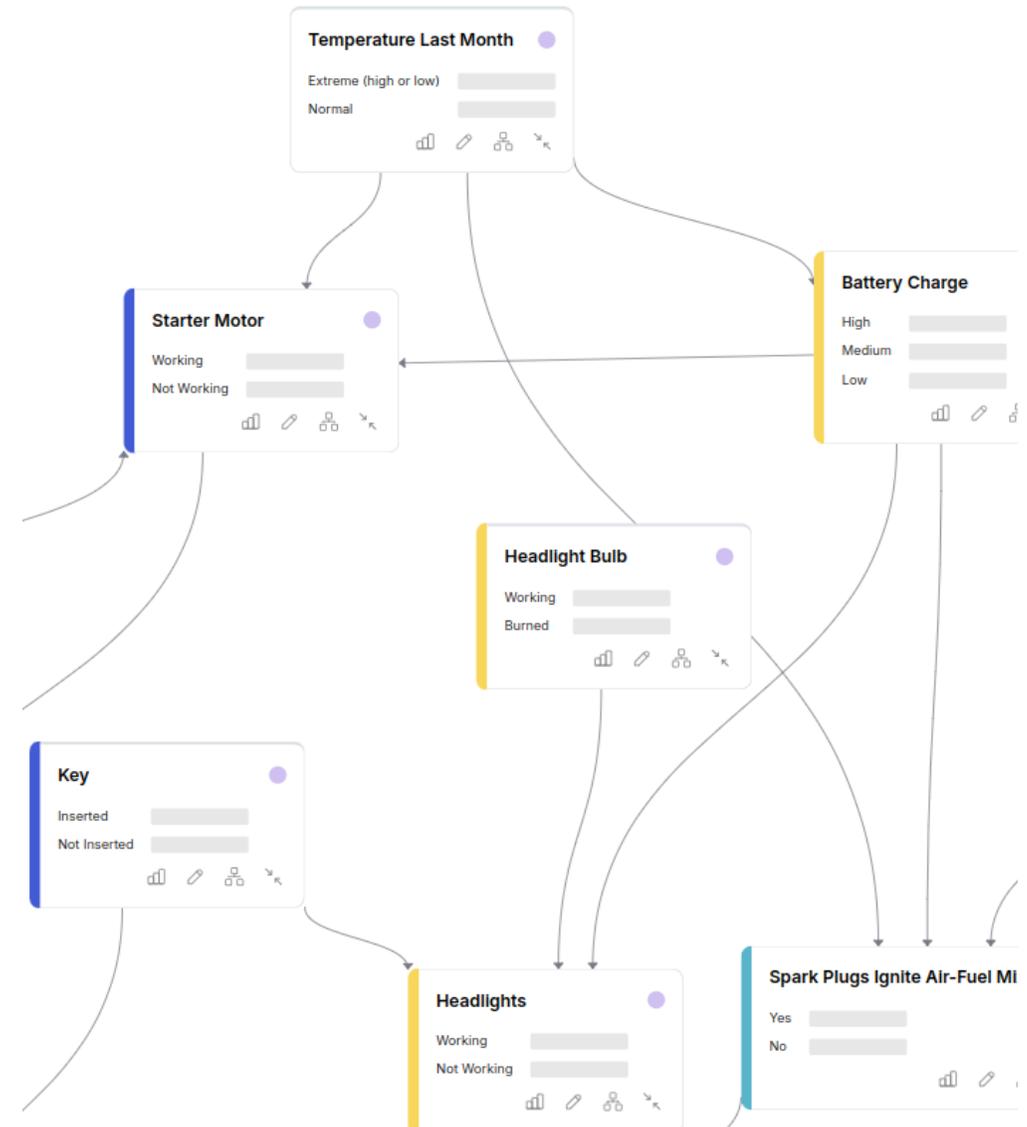
Probability of Y given variable X is artificially set to x_0

Bayesian Networks

Le Reti Bayesiane sono modelli probabilistici grafici che rappresentano le relazioni causali tra variabili in termini probabilistici.

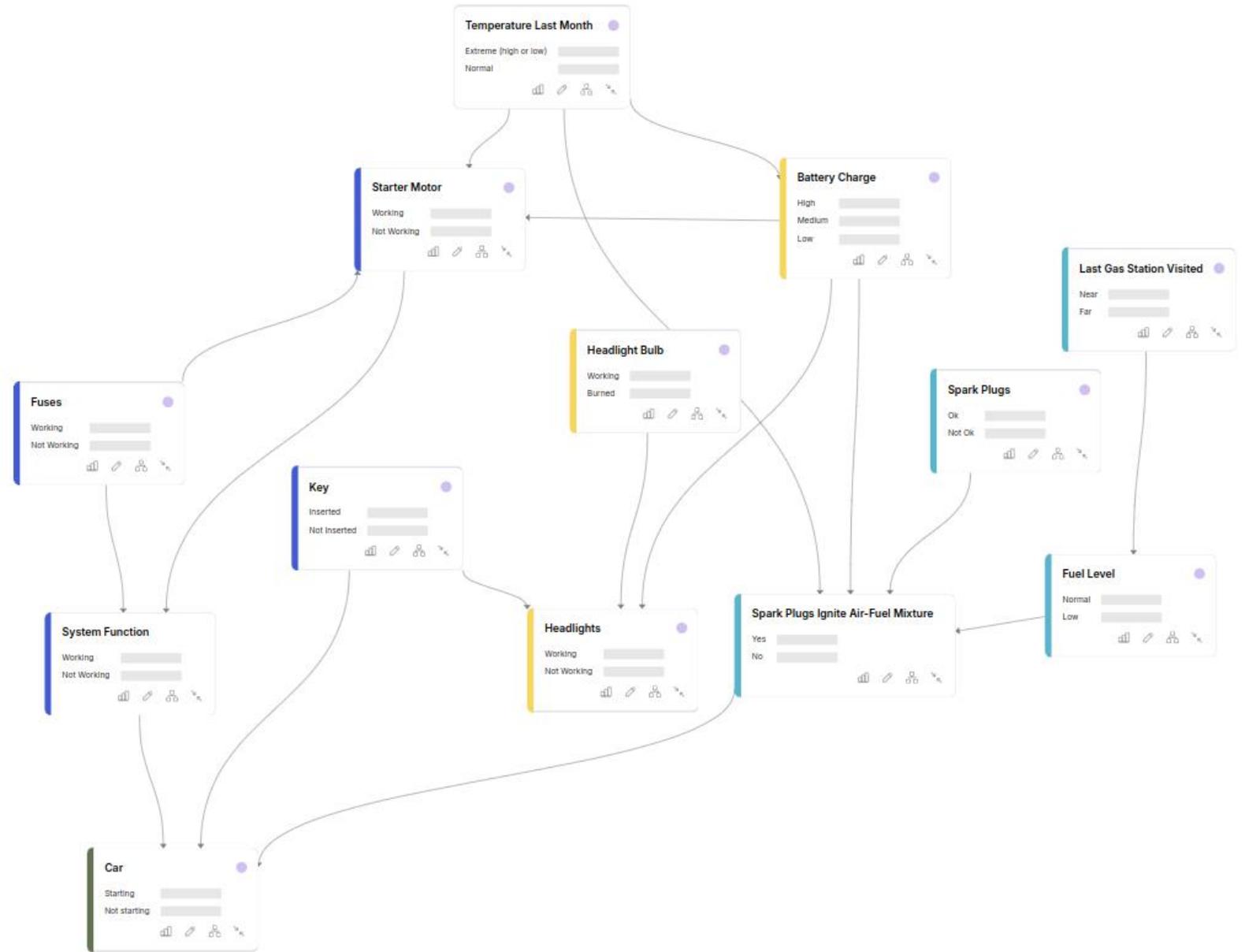
Permettono di codificare:

- Probabilità di accadimento ed incertezza
- Le relazioni tra variabili/informazioni in modo matematicamente appropriato



Esempio:

Accensione macchina



• Terminologia dei grafi

Nodi: unità informativa

Frecce (Edge): collegamenti direzionati che partono da una causa e la collegano alla sua conseguenza

Parenti: come in un albero genealogico, è insieme dei nodi causa di un nodo

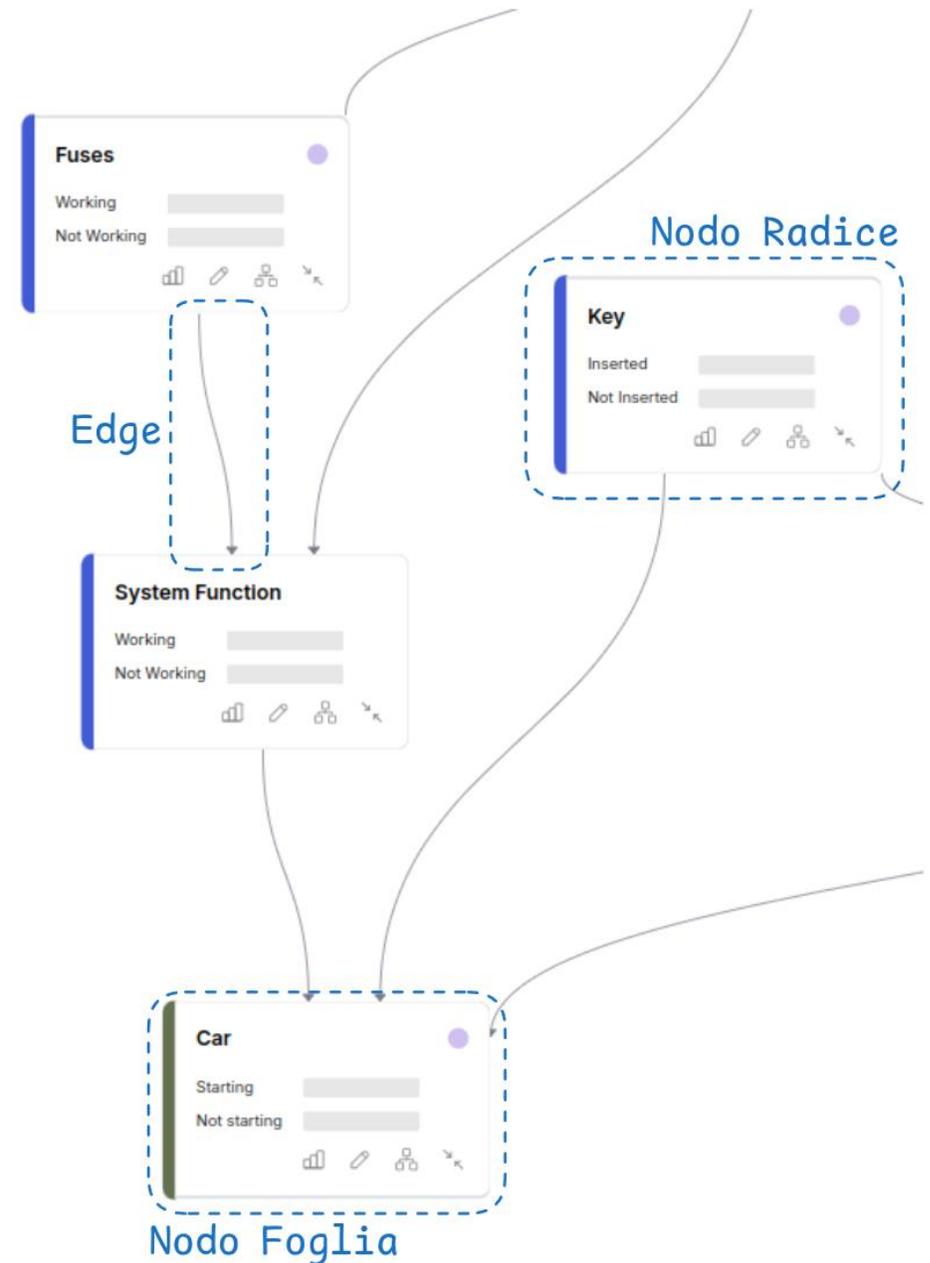
(System Function e Key sono parenti di Car)

Figli: come in un albero genealogico, è l'insieme dei nodi conseguenza di un nodo

(Car è figlio di Key e di System Function)

Nodi foglia: nodi senza alcun figlio

Nodi radice: nodi senza alcun parente

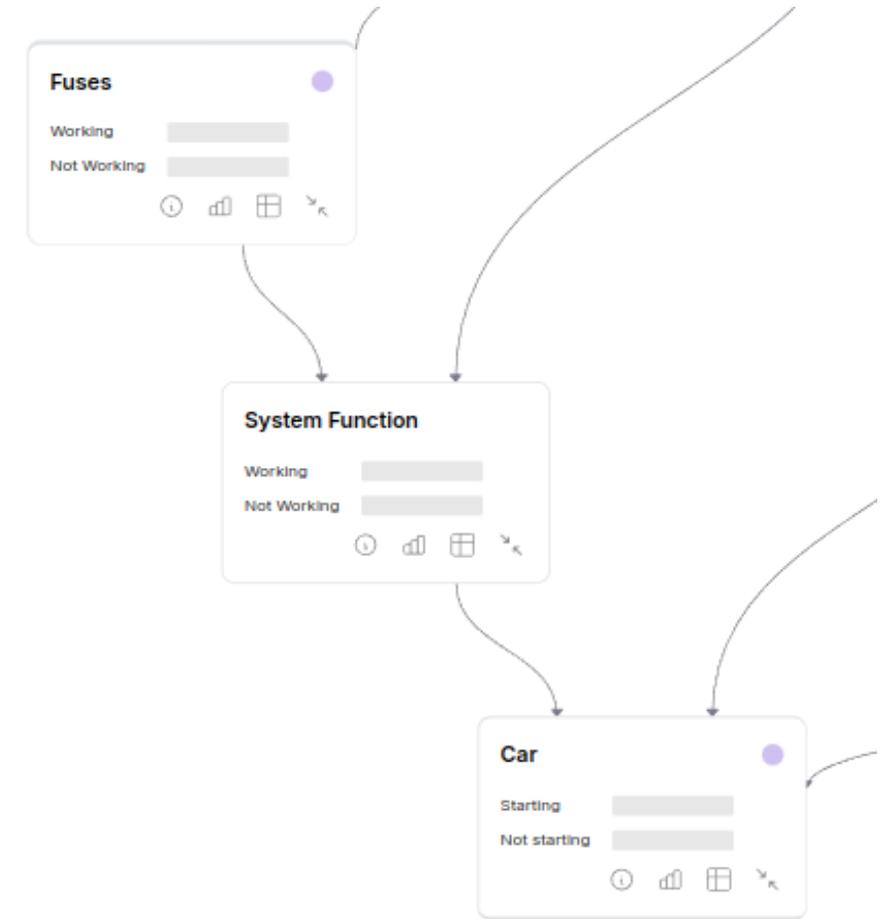


• Cosa codifica un nodo

Concettualmente i nodi rappresentano unità di informazioni all'interno del grafo, che esse siano empiriche quindi con un determinato livello di osservabilità o concettuali ovvero non osservabili ma di puro ragionamento.

Nel percorso che collega una causa alla sua conseguenza i nodi rappresentano tutti gli step dai quali si deve passare.

 *Lo stato dei fusibili implica direttamente che il sistema risulti funzionante o meno e, in ultima analisi, che l'auto si accenda.*



• Nodi Discreti

I nodi sono composti da due elementi principali che ne determinano la validità strutturale: il nome e gli stati.

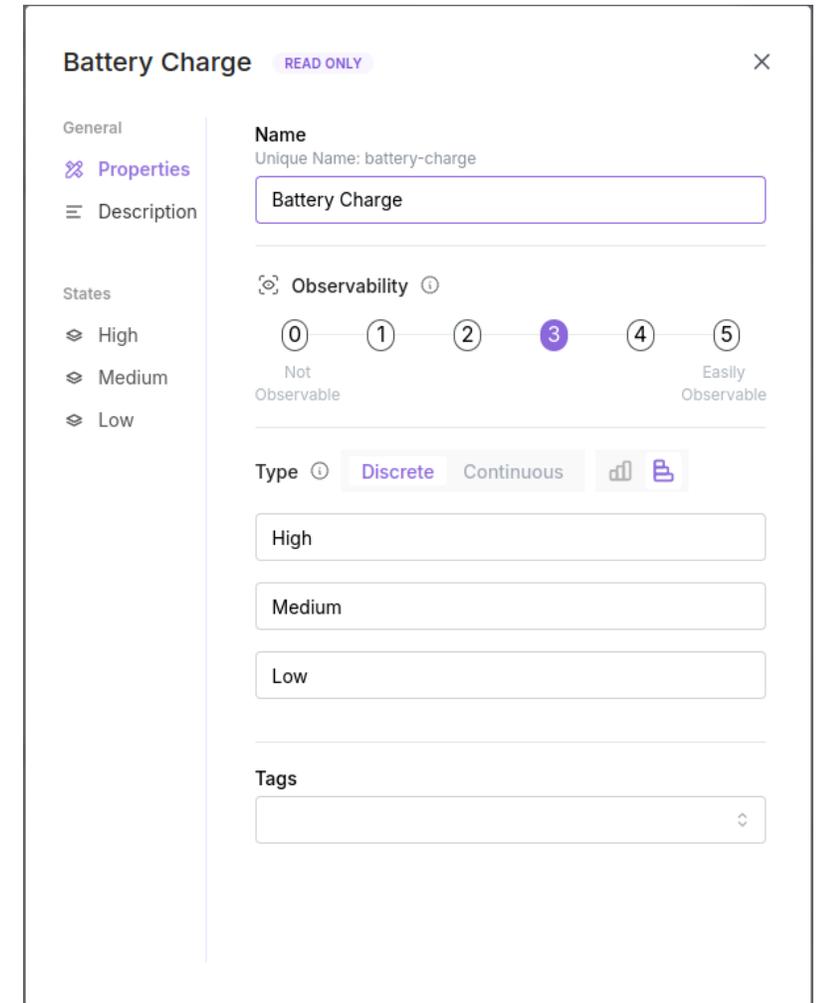
Il nome identifica l'informazione codificata nel nodo in questione mentre gli stati codificano, appunto, i vari stati in cui si può trovare.

 *La carica della batteria (nome) può essere alta, media o bassa (stati)*

Queste informazioni devono essere di carattere generale, non vanno inseriti valori numerici bensì una discretizzazione di questi:

alto – medio – basso, oppure corretto – non corretto.

I dettagli specifici vanno inclusi nelle descrizioni del nodo e dei singoli stati.



Battery Charge READ ONLY

General

✘ Properties

☰ Description

States

High

Medium

Low

Name
Unique Name: battery-charge
Battery Charge

🔗 Observability ⓘ

0 1 2 3 4 5
Not Easily
Observable Observable

Type ⓘ Discrete Continuous

High

Medium

Low

Tags

• Bayesian Network: CPT

La tabella di probabilità condizionata (cpt) codifica la probabilità che il nodo assuma quel valore/stato in funzione di come vengono osservati i suoi parenti.

La probabilità va compilata considerando il contesto di corretto funzionamento non quello problematico.

Questo garantisce manutenibilità del grafo all'aggiunta di nuove problematiche e di informazioni aggiuntive.



Fuel Level

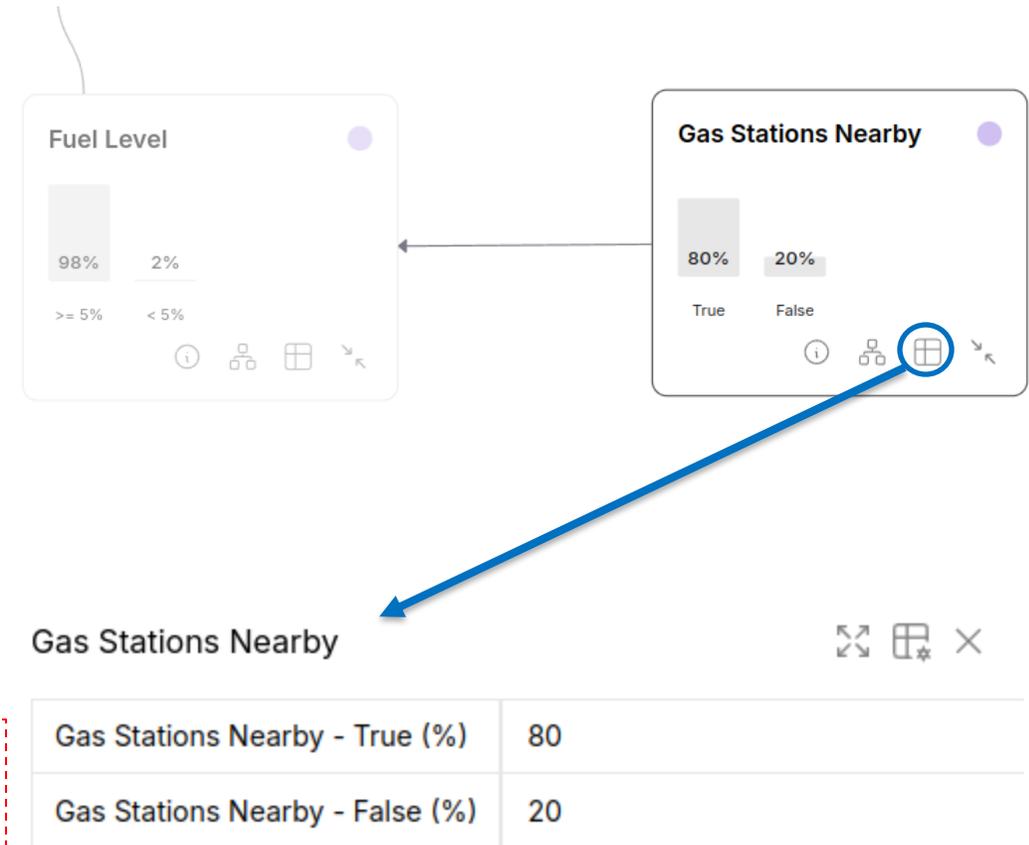
Last Gas Station Visited	Near	Far
Fuel Level - Normal (%)	99	94
Fuel Level - Low (%)	1	6

 "Considerando che l'auto si sia accesa normalmente qual è la probabilità che ..."

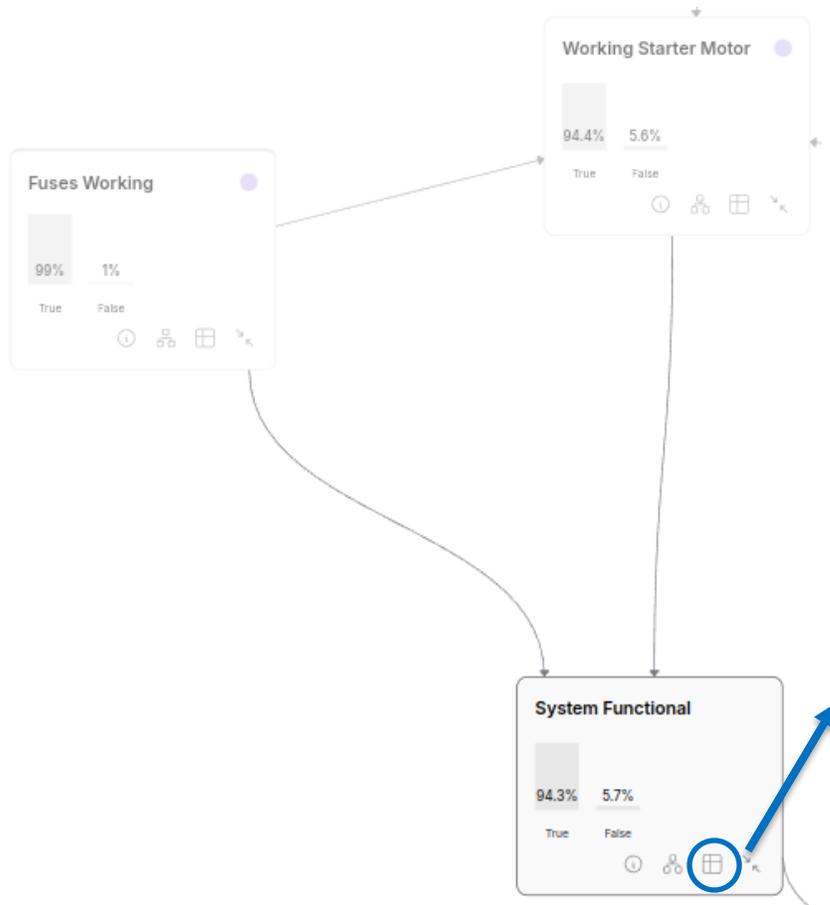
• Bayesian Network: CPT radice

Considerando i nodi radice ovvero quelli senza alcun parente, e quindi senza cause, il ragionamento che sta dietro alla compilazione della tabella è cristallino: bisogna chiedersi quanto spesso capita che quell'informazione assuma quello stato.

 "Quante volte è capitato che *la pompa di benzina (nodo) fosse vicina (stato)*"



• Bayesian Network: CPT figli



Per i nodi che hanno dei genitori e quindi delle cause dirette, la probabilità deve tener conto dello stato in cui si trovano suddette cause. In questo caso il ragionamento alla base della compilazione delle tabelle diventa:

System Functional

Fuses Working	True	True	False	False
Working Starter Motor	True	False	True	False
System Functional - True (%)	99.9	0.1	0.1	0.1
System Functional - False (%)	0.1	99.9	99.9	99.9



"Se lo stato dei fusibili (causa 1) è funzionante e lo starter del motore (causa 2) è funzionante, quanto mi aspetto che anche il sistema generale (nodo da compilare) funzioni?"