



DIRITTO PENALE CONTEMPORANEO

DIRITTO PENALE
CONTEMPORANEO

Fascicolo
11/2018

DIRETTORE RESPONSABILE Gian Luigi Gatta
VICE DIRETTORI Guglielmo Leo, Luca Luparia

ISSN 2039-1676

COMITATO DI DIREZIONE Alexander Bell, Antonio Gullo, Luca Masera, Melissa Miedico, Alfio Valsecchi

REDAZIONE Anna Liscidini (coordinatore), Alberto Aimi, Enrico Andolfatto, Enrico Basile, Carlo Bray, Alessandra Galluccio, Stefano Finocchiaro, Francesco Lazzeri, Erisa Pirgu, Serena Santini, Tommaso Trincherà, Maria Chiara Ubiali, Stefano Zirulia

COMITATO SCIENTIFICO Emilio Dolcini, Novella Galantini, Alberto Alessandri, Jaume Alonso-Cuevillas, Giuseppe Amarelli, Ennio Amodio, Francesco Angioni, Roberto Bartoli, Fabio Basile, Hervé Belluta, Alessandro Bernardi, David Brunelli, Silvia Buzzelli, Alberto Cadoppi, Michele Caianiello, Lucio Camaldo, Stefano Canestrari, Francesco Caprioli, David Carpio, Elena Maria Catalano, Mauro Catenacci, Massimo Ceresa Gastaldo, Mario Chiavario, Luis Chiesa, Cristiano Cupelli, Angela Della Bella, Gian Paolo Demuro, Ombretta Di Giovine, Massimo Donini, Giovanni Fiandaca, Roberto Flor, Luigi Foffani, Gabriele Fornasari, Loredana Garlati, Mitja Gialuz, Glauco Giostra, Giovanni Grasso, Giulio Illuminati, Roberto E. Kostoris, Sergio Lorusso, Stefano Manacorda, Vittorio Manes, Luca Marafioti, Enrico Marzaduri, Jean Pierre Matus, Anna Maria Maugeri, Oliviero Mazza, Alessandro Melchionda, Chantal Meloni, Vincenzo Militello, Santiago Mir Puig, Vincenzo Mongillo, Adan Nieto Martin, Francesco Mucciarelli, Renzo Orlandi, Íñigo Ortiz de Urbina, Francesco Palazzo, Claudia Pecorella, Marco Pelissero, Vicente Pérez-Daudí, Daniela Piana, Lorenzo Picotti, Paolo Pisa, Daniele Piva, Oreste Pollicino, Domenico Pulitanò, Joan Josep Queralt, Tommaso Rafaraci, Paolo Renon, Mario Romano, Gioacchino Romeo, Carlo Ruga Riva, Markus Rübenstahl, Francesca Ruggieri, Marco Scoletta, Sergio Seminara, Rosaria Sicurella, Placido Siracusano, Carlo Sotis, Giulio Ubertis, Antonio Vallini, Paolo Veneziani, Francesco Viganò, Costantino Visconti, Matteo Vizzardi, Francesco Zacchè

Diritto Penale Contemporaneo è un periodico on line, ad accesso libero e senza fine di profitto, nato da un'iniziativa comune di Luca Santa Maria, che ha ideato e finanziato l'iniziativa, e di Francesco Viganò, che ne è stato sin dalle origini il direttore nell'ambito di una partnership che ha coinvolto i docenti, ricercatori e giovani cultori della Sezione di Scienze penalistiche del Dipartimento "C. Beccaria" dell'Università degli Studi di Milano. Attualmente la rivista è edita dall'Associazione "Diritto penale contemporaneo", il cui presidente è l'Avv. Santa Maria e il cui direttore scientifico è il Prof. Gian Luigi Gatta. La direzione, la redazione e il comitato scientifico della rivista coinvolgono oggi docenti e ricercatori di numerose altre università italiane e straniere, nonché autorevoli magistrati ed esponenti del foro.

Tutte le collaborazioni organizzative ed editoriali sono a titolo gratuito e agli autori non sono imposti costi di elaborazione e pubblicazione.

Le opere pubblicate su "Diritto penale contemporaneo" sono attribuite dagli autori con licenza *Creative Commons* "Attribuzione – Non commerciale 3.0" Italia (CC BY-NC 3.0 IT). Sono fatte salve, per gli aspetti non espressamente regolati da tale licenza, le garanzie previste dalla disciplina in tema di protezione del diritto d'autore e di altri diritti connessi al suo esercizio (l. n. 633/1941).

Il lettore può condividere, riprodurre, distribuire, stampare, comunicare al pubblico, esporre in pubblico, cercare e segnalare tramite collegamento ipertestuale ogni lavoro pubblicato su "Diritto penale contemporaneo", con qualsiasi mezzo e formato, per qualsiasi scopo lecito e non commerciale, nei limiti consentiti dalla licenza *Creative Commons* "Attribuzione – Non commerciale 3.0 Italia" (CC BY-NC 3.0 IT), in particolare conservando l'indicazione della fonte, del logo e del formato grafico originale, nonché dell'autore del contributo.

La rivista fa proprio il Code of Conduct and Best Practice Guidelines for Journal Editors elaborato dal COPE (Committee on Publication Ethics).

Peer review.

Salvo che sia diversamente indicato, tutti i contributi pubblicati nella sezione *papers* di questo fascicolo hanno superato una procedura di *peer review*, attuata secondo principi di trasparenza, autonomia e indiscusso prestigio scientifico dei revisori, individuati secondo criteri di competenza tematica e di rotazione all'interno dei membri del Comitato scientifico. Ciascun lavoro soggetto alla procedura viene esaminato in forma anonima da un revisore, il quale esprime il suo parere in forma parimenti anonima sulla conformità del lavoro agli standard qualitativi delle migliori riviste di settore. La pubblicazione del lavoro presuppone il parere favorevole del revisore. Di tutte le operazioni compiute nella procedura di *peer review* è conservata idonea documentazione presso la redazione.

Modalità di citazione.

Per la citazione dei contributi presenti nei fascicoli di *Diritto penale contemporaneo*, si consiglia di utilizzare la forma di seguito esemplificata: N. COGNOME, *Titolo del contributo*, in *Dir. pen. cont.*, fasc. 1/2017, p. 5 ss.

PROVA GENETICA DEL DNA E RISULTATI DISSONANTI: COME VALUTARE CONGIUNTAMENTE GLI ELEMENTI SCIENTIFICI DI PROVA

di Franco Taroni, Ilaria De March, Paolo Garbolino e Silvia Bozza

SOMMARIO: 1. Introduzione. – 2. Le sentenze delle Corti di Assise e della Corte Suprema di Cassazione. – 3. Il problema scientifico e giuridico: il valore congiunto degli elementi di prova raccolti. – 4. I modelli grafici probabilistici: le Reti Bayesiane. – 4.1. Elementi fondamentali delle Reti Bayesiane. – 4.2. Un esempio semplice di applicazione delle Reti Bayesiane. – 5. L'analisi congiunta delle evidenze: conflitto e soluzione. – 6. Conclusione.

1. Introduzione.

Lo scopo di questo articolo è condurre una riflessione in merito al valore da attribuire a risultati dissonanti¹ provenienti da due analisi del DNA e specificatamente sulla discordanza tra, da una parte, il risultato genetico negativo ottenuto a partire dal campionamento effettuato sul reperto principale in un procedimento penale – una collana in metallo – dal perito di un laboratorio di Genetica Forense e, dall'altra, il risultato positivo ottenuto successivamente sullo stesso reperto dal perito di un secondo laboratorio di Genetica Forense, che ha messo in evidenza la presenza di un profilo genetico corrispondente a quello della vittima. Quest'ultimo risultato è stato considerato altamente incriminante da parte dell'Accusa nei confronti della persona sospettata, nel cui appartamento è stata rinvenuta la collana.

Questo lavoro segue un nostro precedente articolo pubblicato in questa *Rivista*², dove, dopo una breve esposizione del caso in esame – l'aggressione e l'omicidio di un'anziana signora (che successivamente indicheremo come 'la vittima'), il rinvenimento di una medaglietta sprovvista di catenina che la vittima portava attorno al collo e il ritrovamento di una catenina in oro spezzata nell'abitazione della persona sospettata del crimine – sono stati analizzati i documenti relativi al fascicolo, specificatamente i rilievi tecnici effettuati nell'appartamento della vittima e

¹ Si usa il termine dissonante per indicare che i risultati di due analisi di laboratorio sullo stesso reperto puntano in direzioni diverse. Si veda D.A. Schum, *Evidential foundations of probabilistic reasoning*, John Wiley & Sons, New York 1994, per una discussione approfondita di tali situazioni.

² F. TARONI, S. BOZZA, P. GARBOLINO, [Contaminazioni di un reperto con il DNA. Quando la prova genetica porta direttamente alla condanna](#), in questa *Rivista*, 12 febbraio 2018.



11/2018

nell'abitazione della persona sospettata, le relazioni tecniche di genetica forense redatte dai periti e le annotazioni inerenti le attività compiute sui reperti analizzati, e si è prodotta una valutazione in merito alle due relazioni tecniche di genetica forense provenienti dai due laboratori. In particolare, partendo dalla lettura di questi documenti, non è stato possibile escludere che il profilo genetico trovato sulla catenina in metallo ritrovata nell'abitazione della persona sospettata e dichiarato corrispondente al profilo della vittima, come messo in evidenza nella relazione tecnica del secondo perito, non sia stato ottenuto in conseguenza a una contaminazione del reperto. In effetti, l'analisi condotta aveva evidenziato alcune criticità (tra le quali, per esempio, la cronologia delle attività svolte presso il secondo laboratorio, l'assenza di quantificazione delle potenziali contaminazioni commesse nei test di controllo di qualità) che andrebbero a sostegno di una possibile contaminazione.

Mentre nel precedente lavoro ci si è soffermati a sottolineare perché l'esistenza di una contaminazione a spiegazione del risultato positivo ottenuto dal secondo laboratorio non possa essere esclusa, l'obiettivo del presente articolo è di approfondire come sia possibile valutare congiuntamente risultati dissonanti provenienti da distinti esami genetici.

Questo permetterà inoltre di giustificare la conclusione del nostro primo articolo dove si scriveva: «La nostra analisi permette di concludere che il valore da accordare ai risultati del secondo laboratorio è neutro (se non, addirittura, in favore della Difesa); ciò significa che l'osservazione del profilo genetico dichiarato corrispondente con quello della vittima e trovato sulla catenina di metallo in scarse quantità non permette di discriminare tra le ipotesi di interesse per le parti. I risultati delle analisi genetiche non permettono né di sostenere l'ipotesi che la persona sospettata abbia strappato la collana alla vittima, né di sostenere l'ipotesi che la persona sospettata non abbia strappato la collana alla vittima».

2. Le sentenze delle Corti di Assise e della Corte Suprema di Cassazione.

Sia nella sentenza di primo grado della Corte di Assise, che in quella della Corte di Assise d'Appello, i giudici si sono posti la domanda di come si dovrebbe reagire davanti a due risultati dissonanti al fine di gestirli in modo congruo. I giudici di entrambe le Corti hanno raggiunto le loro conclusioni attribuendo credibilità esclusivamente al secondo laboratorio (quello che ha messo in evidenza una corrispondenza genetica tra il DNA osservato sulla catenina e quello della vittima) a discapito del primo. Nella sentenza di primo grado si può leggere:

«Concludendo sul punto, si ribadisce che sulla collana non vi erano tracce di DNA riferentisi alle sorelle [nome persona sospettata], che quelle relative alla [nome vittima], rilevate da [nome perito del 2° laboratorio], non derivano da una supposta contaminazione tra i reperti, ma dalla diversa metodica usata da quest'ultima: in particolare la procedura per immersione e la diversità di kit utilizzato ha consentito di rilevare materiale rimasto

impigliato tra maglia e maglia, che il metodo di tamponamento usato da [nome perito del 1° laboratorio] non aveva invece potuto rilevare. » (p. 7),

mentre nella sentenza di appello si afferma:

«[...] che la presenza di esso non è giustificabile o altrimenti spiegabile con una contaminazione fra campioni e reperti; che la differenza fra gli esiti delle analisi effettuate dal [nomi dei due periti] è una diretta conseguenza delle diverse metodiche utilizzate rispettivamente dalle due consulenti. Infatti, mentre la prima ha effettuato la ricerca con il metodo del tamponamento, la seconda, seguendo la metodica dell'immersione, ha rinvenuto, impigliato fra una maglia e l'altra della collana, la traccia di materiale biologico che pur minima era perfettamente integra ed ha consentito di estrapolarne il DNA della vittima» (p. 24).

Secondo quanto concluso dalle due Corti, solo il risultato del secondo laboratorio sarebbe attendibile, in virtù dell'applicazione di una metodologia di campionamento ed estrazione considerata, dalle Corti stesse, più efficace rispetto a quella impiegata nel primo laboratorio³.

I giudici della Corte Suprema di Cassazione hanno ribadito come la differenza tra i risultati analitici sul reperto sia prettamente riconducibile alle diverse metodiche applicate dai due laboratori per l'estrazione di materiale biologico potenzialmente presente sul reperto, arrivando pertanto alla conclusione che la potenziale contaminazione del reperto possa essere esclusa. La Corte Suprema di Cassazione ha dunque accolto gli argomenti sostenuti dai giudici delle corti di prima e seconda istanza, scrivendo:

«Osserva il Collegio che su tali doglianze, che ricalcano in gran parte quelle già formulate dalla difesa con i motivi di appello, il provvedimento ha fornito ampia e puntuale risposta con motivazione adeguata, logica e non contraddittoria. In particolare la sentenza di secondo grado - dopo aver posto in rilievo che la stessa difesa non pone in discussione l'esattezza del risultato dell'esame del DNA estratto dalla traccia e la sua riconducibilità alla

³ Omettiamo dalla nostra analisi una discussione sulla fallacia commessa dalla seconda corte che afferma che il DNA osservato è quello della vittima quando invece il laboratorio ha unicamente messo in evidenza un profilo DNA corrispondente a quello della vittima. Il fatto di osservare un profilo genetico identico non implica logicamente che questo profilo sia quello della vittima. Il lettore interessato può fare riferimento, per esempio, a W.C. THOMPSON, E.L. SCHUMANN, *Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defence attorney's fallacy*, in *Law and Human Behaviour*, vol. 11, 1987, p. 167-187; I.W. EVETT, *Avoiding the transposed conditional*, in *Science & Justice*, vol. 35, 1995, p. 127-131; C. AITKEN, F. TARONI, *Statistics and the evaluation of evidence for forensic scientist*, 2nd Ed., John Wiley & Sons, Chichester, 2004, per un'analisi attenta di questo errore logico conosciuto con i nomi inglesi di *prosecutor's fallacy* (errore del procuratore) o di *transposed conditional* (trasposizione del condizionale), questo perché l'errore si annida nella confusione tra due probabilità condizionate.



11/2018

vittima, ma prospetta l'inattendibilità di quel risultato perché, in ipotesi, frutto di una contaminazione del reperto su cui la traccia è stata rinvenuta - illustra i motivi per i quali esclude l'ipotizzata contaminazione.

[...] il provvedimento, inoltre, fornisce adeguata risposta anche in merito alle obiezioni difensive sul diverso esito del risultato degli esami effettuati sul monile, chiarendo che non sussiste, contrariamente a quanto asserito nel ricorso, alcuna contrapposizione di risultati scientifici, ma che la differenza dell'esito del controllo effettuato sulla catenina in sequestro, è una diretta conseguenza delle diverse tecniche di analisi utilizzate dai due consulenti; in particolare, il metodo del tamponamento, utilizzato dal primo consulente e quello dell'immersione, utilizzato dal secondo, che ha permesso di rinvenire "impigliata fra una maglia e l'altra della collana", una traccia di materiale biologico, sfuggita alla prima tecnica di rilevamento. In merito all'ipotizzata eccessiva esiguità della traccia, la sentenza, precisa che, seppur minima, la traccia rilevata era perfettamente integra, tanto da consentire l'estrapolazione del DNA della vittima, circostanza quest'ultima, ribadisce il giudice di appello, non contestata neanche dalla difesa dell'imputata. Inoltre il provvedimento chiarisce che la circostanza che la traccia fosse minima non contrasta per niente, come sostenuto dalla difesa, con il fatto che la collanina fosse abitualmente indossata dalla vittima, potendosi, al contrario, affermare che, proprio tale aspetto, conferma la circostanza che quella collanina fu accuratamente pulita da chi la custodì dopo il delitto» (pp. 10-11).

È dunque consolidato, a risoluzione del conflitto tra le due perizie, il favore dei giudici in accoglimento dei risultati genetici ottenuti dal secondo laboratorio, a discapito di quelli emersi dalle analisi svolte in precedenza dal primo laboratorio. Ma allora, cosa si dovrebbe fare dell'informazione fornita dal primo laboratorio? Semplicemente escluderla? Noi non siamo di questo avviso.

Prima di illustrare il nostro punto di vista, che si discosta da quanto concluso dalle diverse Corti, dobbiamo chiarire un punto sottolineato da tutte e tre le Corti. La difesa non pone in alcun modo in discussione l'affidabilità del risultato dell'esame del DNA estratto dalla traccia e la sua riconducibilità alla vittima. Non viene infatti criticato il risultato dell'analisi: il DNA osservato presenta veramente il profilo messo in evidenza dal laboratorio e questo profilo è identico a quello della vittima; la difesa non contesta infatti che si possa affermare che questo profilo corrisponda a quello della vittima. Il problema sollevato dalla difesa deriva dal fatto che vi sono due risultati dissonanti e che il risultato positivo del secondo laboratorio può essere riconducibile a una contaminazione con sangue della vittima.

3. Il problema scientifico e giuridico: il valore congiunto degli elementi di prova raccolti.

La domanda che ci poniamo e alla quale cercheremo di dare una risposta logica è la seguente: dobbiamo veramente scegliere – giustificando ovviamente la nostra scelta – uno solo tra i due risultati, o dovremmo invece valutare entrambi i risultati ottenuti tenendo conto delle diverse metodologie applicate dai laboratori e giungere così a una valutazione congiunta degli elementi di prova?

La letteratura scientifica, giuridica e psicologica ha, a più riprese, sottolineato l'utilità di metodi per l'analisi formale del ragionamento razionale. Un esempio è offerto dal libro del premio Nobel Daniel Kahnemann⁴, dove si sottolinea che la gestione di sistemi complessi di circostanze – dunque situazioni caratterizzate da molteplici variabili che potenzialmente presentano legami di dipendenza fra di loro – necessita di un'assistenza logica. Metodi di ragionamento formale devono essere capaci di assistere sia gli scienziati forensi che i giudici nella gestione dell'informazione complessa. Per esempio, la comprensione dei legami di dipendenza che possono esistere tra diverse osservazioni dev'essere gestita in modo ottimale. Focalizzarsi unicamente sulla valutazione di singole evidenze, ignorando il contesto nel quale sono state osservate, rappresenta una limitazione inaccettabile, come spesso sottolineato nella letteratura scientifica, idonea ad alimentare conclusioni errate⁵. Ad esempio, risposte alle domande del tipo 'i due elementi di prova provocano della ridondanza o piuttosto degli effetti sinergici?'⁶ rappresentano informazioni importanti e necessarie per una corretta valutazione degli elementi di prova in un determinato contesto. Lo statistico Dennis Lindley⁷ pone grande enfasi sulla complessità del fenomeno di valutazione congiunta degli elementi di prova (o evidenze). La letteratura giuridica sottolinea ugualmente l'attuale incapacità delle scienze forensi di rispondere a domande fondamentali, una delle quali riguarda la valutazione congiunta di elementi di prova⁸, ed è indubbio che nella maggior parte delle attività inferenziali condotte regolarmente nella vita quotidiana ci si trovi o ci si venga a trovare di fronte a molteplici informazioni che necessitano d'essere valutate in modo congruo. È dunque nostro compito investigare se

⁴ D. KAHNEMANN, *Pensieri lenti e veloci*, Mondadori, Milano 2012.

⁵ P.O. JUCHLI, *Combining evidence*. Ph.D dissertation, University of Lausanne, 2016.

⁶ Per una discussione approfondita su questi aspetti si veda D.A. SCHUM, *Evidential foundations of probabilistic reasoning*, John Wiley & Sons, New York 1994.

⁷ D.V. LINDLEY, *Foreword*, in C. Aitken, F. Taroni, *Statistics and the evaluation of evidence for forensic scientists*. 2nd ed., John Wiley & Sons, Chichester 2004, p. XXII-XXIV, scrive: « [...] probability is designed to effect such combinations but the accumulation of simple rules can produce complicated procedures».

⁸ M. REDMAYNE, *Quality and forensic science evidence: an overview*, in J.F. Nijboe, W.J.J.M. Sprangers (eds.), *Harmonisation in Forensic Expertise: An inquiry into desirability of and opportunities for international standards*. Thela Thesis, Amsterdam, 2000, p. 299-317. « [...] a theoretical stagnation of forensic science. The science's practical bent rarely confronts the fundamental question » (p. 305). Si veda anche D.H. KAYE, *Beyond uniqueness: the birthday paradox, source attribution and individualization in forensic science testimony*, in *Law, Probability and Risk*, vol. 12, 2013, p. 3-11.

vi siano modi di descrivere le relazioni tra molteplici elementi di prova e come questi si interfacciano l'uno con l'altro.

Una possibile soluzione è offerta dai modelli grafici probabilistici, che offrono il supporto metodologico necessario per i) rappresentare schematicamente tutte le variabili in gioco e le relazioni tra esse, e ii) calcolare l'impatto di una e più osservazioni (evidenze) sulle ipotesi d'interesse per il caso in esame. Per esempio, se fossimo interessati alla probabilità dell'ipotesi che la collana sia stata strappata alla vittima dalla persona sospettata, come potremmo aggiornare, e come cambierebbe, la probabilità iniziale (in assenza di riscontri di laboratorio) alla luce di tutti gli elementi (evidenze) osservati?

4. I modelli grafici probabilistici: le Reti Bayesiane.

Le Reti Bayesiane, in inglese, *Bayesian Networks*, termine abbreviato con le lettere BNs, possono essere genericamente definite come una rappresentazione schematica delle relazioni (rappresentate graficamente da frecce) esistenti tra variabili (rappresentate da nodi sotto forma grafica di ovali) considerate rilevanti per risolvere lo specifico problema inferenziale (si veda, a puro titolo illustrativo, la Rete Bayesiana rappresentata nella *Figura 1*).

A partire dagli inizi degli anni '80, i BNs hanno ricevuto un'attenzione vieppiù crescente in diversi settori scientifici operanti nella gestione della conoscenza. La loro diffusione si spiega con l'abilità di questi modelli nel coordinare il ragionamento probabilistico alla base delle inferenze induttive che costituiscono il problema oggetto di interesse: com'è possibile aggiornare lo stato di incertezza a mano a mano che nuove informazioni si rendono disponibili?

Le Reti Bayesiane costituiscono un valido strumento per la rappresentazione dell'incertezza, per la sua gestione, e possono pertanto essere preziose per dare una risposta razionale alla precedente domanda. L'aggettivo 'Bayesiana', in associazione al sostantivo 'Rete', deriva dal Teorema di Bayes⁹, un formidabile risultato della teoria del calcolo delle probabilità che offre una soluzione elegante e non controversa per aggiornare lo stato d'incertezza a mano a mano che nuove informazioni si rendono disponibili.

4.1. Elementi fondamentali delle Reti Bayesiane.

Le Reti Bayesiane rappresentano una combinazione della teoria dei grafi, che consente di riprodurre la struttura qualitativa del problema oggetto di interesse (nel seguito, chiamata modello), e della teoria delle probabilità, che viene utilizzata per

⁹ Si veda, per esempio, P. GARBOLINO, *Probabilità e logica della prova*, Giuffrè Editore, Milano 2014; F. TARONI, S. BOZZA, J. VUILLE, *Il ruolo della probabilità nella valutazione della prova scientifica*, in G. Canzio, L. Luparia (a cura di), *Prova scientifica e processo penale*, Cedam, Milano 2018, p. 23-74.



11/2018

caratterizzare la natura e la forza dei legami presenti nel modello così proposto. In modo più formale, una Rete Bayesiana è composta dai seguenti elementi:

- Una lista di variabili rappresentate graficamente da nodi. Ognuno di questi nodi possiede una lista finita di stati tra loro esclusivi (uno solo può essere vero) ed esaustivi (la lista comprende tutti gli stati possibili).
- Una serie di frecce che connettono tra loro i nodi. Le frecce indicano la dipendenza probabilistica tra i nodi (una freccia da un nodo A ad un nodo B indica che A ha influenza su B, o viceversa, che B dipende da A). La lista delle variabili e la serie di frecce si combinano in modo da ottenere un grafo detto 'aciclico' dove l'esistenza di cicli (o *loops*) non è permessa¹⁰.
- A ciascuna variabile (nodo) della Rete è associata una tabella di probabilità contenente le probabilità attribuite a ciascuno stato della variabile. La tabella di probabilità associata ad una variabile A con due stati (ad esempio, 'vero' e 'falso') e che non presenti frecce provenienti da altre variabili, conterrà la probabilità (non condizionata) $\Pr(A)$, nello specifico $\Pr(A=\text{'vero'})$ e $\Pr(A=\text{'falso'})$. Il termine 'non condizionata' si riferisce unicamente all'assenza di legami probabilistici tra la variabile A e le altre variabili. Ovviamente, la probabilità $\Pr(A)$ è da considerare condizionata al contesto generale e alle informazioni disponibili. Per esempio, se fossimo interessati alla probabilità che piova oggi, la nostra valutazione dovrebbe essere condizionata, ad esempio, al luogo. La probabilità che piova oggi a Palermo non sarà certo la stessa della probabilità che piova oggi a Edimburgo. Il contesto, in questo caso geografico, condiziona la nostra valutazione.

Nel caso in cui la variabile A dipenda da più variabili, per esempio B_1, \dots, B_n , (e dunque vi siano più frecce che puntano dalle variabili B_1, \dots, B_n sulla variabile A), la tabella di probabilità associata alla variabile A sarà più complessa perché dovrà contenere tutte le probabilità condizionate $\Pr(A | B_1, \dots, B_n)$, vale a dire le probabilità di ciascuno stato della variabile A condizionate a ciascun possibile stato delle variabili B_1, \dots, B_n (il condizionamento è espresso dalla barra verticale |).

I nodi di una Rete Bayesiana rappresentano dunque le variabili in gioco, all'interno di un modello di ragionamento ipotetico nel quale, a priori, nessuna variabile può essere assunta come vera, affinché il suo valore probatorio possa essere valutato per mezzo del Teorema di Bayes. Per esempio, quando costruiamo il modello, il fatto che il perito abbia osservato quel particolare profilo genetico è un fatto noto, ma alla variabile che rappresenta la proposizione 'Il perito osserva il profilo genetico X', non dobbiamo attribuire probabilità pari a uno, perché nella costruzione del modello dobbiamo chiederci invece: qual è la probabilità che la proposizione 'Il perito osserva il profilo genetico X' sia vera, se la proposizione 'Il sospetto è all'origine della traccia ritrovata sui luoghi del crimine' è vera? E qual è la probabilità che la stessa proposizione, cioè 'Il perito osserva il profilo genetico X' sia vera, se la proposizione 'Il sospetto è all'origine

¹⁰ Questo significa che non è possibile, partendo da un nodo, e seguendo la direzione delle frecce, tornare al punto di partenza o allo stesso nodo.



11/2018

della traccia ritrovata sui luoghi del crimine' è, invece, falsa? Per fare un altro esempio, il fatto che il Sig. X, testimone, abbia dichiarato che Y è noto quando costruiamo il modello, ma alla proposizione 'Il Sig. X, testimone, dichiara che Y' non dobbiamo attribuire probabilità pari a uno, perché dobbiamo, invece, chiederci: qual è la probabilità che la proposizione 'Il Sig. X, testimone, dichiara che Y' sia vera, se la proposizione 'Y è accaduto' è vera? E qual è la probabilità che la stessa proposizione, cioè 'Il Sig. X, testimone, dichiara che Y' sia vera, se la proposizione 'Y è accaduto' è, invece, falsa?

Ovviamente è indispensabile che a tali affermazioni possano essere attribuite delle probabilità (in funzione delle conoscenze dell'utilizzatore del modello), ad esempio la probabilità che un dato profilo genetico sia osservato dal perito, dato che il soggetto non è all'origine della traccia, o la probabilità che il Sig. X affermi di aver visto il sospetto in un certo luogo a una certa ora, dato che il sospetto era in un certo luogo a una certa ora, ovvero la probabilità che il testimone sia cognitivamente e ideologicamente attendibile. Questo modo di strutturare la conoscenza caratterizza le nostre azioni e comportamenti quotidiani.

Va sottolineato che gli stati delle variabili non sono necessariamente sempre di tipo binario (ad esempio, 'vero' e 'falso', come nel caso ora presentato), ma possono includere liste (per esempio, una lista di colori, {rosso, verde, blu}, nel caso in cui la variabile di interesse sia il colore di un'autovettura rubata), oppure valori (o intervalli di valori) quando la variabile è frutto di una misurazione (per esempio, la lunghezza dei capelli del criminale osservato da dei testimoni).

Gli stati delle variabili possono rappresentare essi stessi delle affermazioni più complesse. Nel nostro caso, le versioni dei fatti proposte dall'accusa ('La collana appartiene alla vittima, ed è stata strappata dalla sospetta durante l'aggressione') e dalla difesa ('La collana appartiene alla sospetta, che non l'ha toccata per anni e non ne ricorda la provenienza') rappresentano gli stati di una variabile che possiamo genericamente chiamare H , e che rappresenta il modo in cui la vicenda si è effettivamente svolta.

Le frecce nella Rete Bayesiana rappresentano, come osservato poc'anzi, delle relazioni probabilistiche ritenute fondate tra le variabili del modello proposto per risolvere un problema inferenziale. Se un modello è costruito correttamente, allora una freccia che parte da un nodo A verso un nodo B indica che la variabile A ha un'influenza diretta sulla variabile B . L'esistenza di molteplici frecce verso una variabile, o il numero di stati della stessa, hanno come conseguenza l'ampliamento delle tabelle di probabilità legate a queste variabili, a completamento delle quali si potranno usare le probabilità dette 'soggettive', idonee ad esprimere il grado di credenza o di fiducia dell'operatore¹¹.

¹¹ Si noti, a fronte di un possibile scetticismo nei confronti di un'impostazione soggettivista delle probabilità, che nulla impedisce che queste probabilità siano alimentate, qualora disponibili, da valori statistici espressi in termini di frequenze relative. Si veda F. TARONI, P. GARBOLINO, A. BIEDERMANN, C. AITKEN, S. BOZZA, *Reconciliation of subjective probabilities and frequencies in forensic science*, in *Law, Probability & Risk*, 2018, in corso di pubblicazione, disponibile al link <https://doi.org/10.1093/lpr/mgy014>.

Le tabelle rappresentano dunque l'interfaccia tra il modello qualitativo (descrizione di nodi e frecce) e i dati numerici che quantificano i legami (frecce) messe in evidenza¹².

4.2. Un esempio semplice di applicazione delle Reti Bayesiane.

Come scritto in precedenza, lo scopo operativo di una Rete Bayesiana è di rappresentare il problema oggetto di interesse in tutte le sue componenti, esplicitando le relazioni tra quest'ultime, e di processare ogni nuova informazione acquisita dall'operatore, aggiornando le tabelle di probabilità condizionate per le variabili che dipendono da questa informazione. In altre parole, la Rete sarà utilizzata per calcolare la probabilità condizionata (alla nuova informazione) di ogni altra variabile descritta dal modello. Per esempio, sarà possibile calcolare la probabilità associata agli stati ('vero' e 'falso') della variabile 'La collana è stata strappata dalla persona accusata dell'omicidio' inserendo lo stato 'vero' della variabile 'Il perito ha osservato il profilo genetico della vittima'.

Generalmente, la prima variabile del nostro esempio è chiamata *variabile ipotesi* perché l'interesse dell'operatore è essenzialmente posto sulla conoscenza di questa variabile. L'osservazione genetica riportata dal perito (la presenza del genotipo della vittima sulla collana) è chiamata *variabile evidenza*.

Lo scopo è dunque quello di calcolare la probabilità degli stati della *variabile ipotesi* dopo che alla *variabile* (o alle *variabili*) *evidenza* sia stato attribuito lo stato 'vero'.

Si tratta dunque di poter rappresentare e aggiornare l'incertezza legata alla variabile di interesse per la Corte. Le Reti Bayesiane, che si appoggiano su algoritmi per il calcolo delle probabilità accettati e validati dalla comunità scientifica, offrono il grande vantaggio di distogliere l'attenzione dai problemi di calcolo, che diventano velocemente ingestibili all'aumentare del numero delle variabili e/o del numero di stati delle stesse, a favore della verifica della correttezza del modello proposto: dalla scelta delle variabili, alla definizione della loro struttura di dipendenza fino alla misurazione probabilistica dell'incertezza ad esse relativa.

Al fine di rendere esplicito questo meccanismo d'apprendimento, consideriamo una semplice rete composta da due nodi binari: il nodo ipotesi H (dal termine inglese *Hypothesis*), e il nodo evidenza E (dal termine inglese *Evidence*). Immaginiamo che il nodo ipotesi faccia riferimento all'ipotesi d'interesse per la Corte, per esempio, 'Il sospetto è all'origine della traccia di origine biologica ritrovata sui luoghi del crimine' e l'evidenza E si riferisca al possibile risultato di laboratorio 'Vi è una corrispondenza tra i profili genetici della traccia e del sospetto'. Gli stati delle due variabili possono essere descritti nei termini di 'vero' e 'falso'.

¹² Dettagli sulle proprietà matematiche delle Reti Bayesiane non saranno presentate in questo articolo. Il lettore interessato potrà far riferimento a F. TARONI, A. BIEDERMANN, S. Bozza, P. Garbolino, C. Aitken, *Bayesian networks for probabilistic inference and decision analysis in forensic science*, 2nd ed., John Wiley & Sons, Chichester 2014, e ai riferimenti bibliografici citati nell'opera.

La tappa successiva nella costruzione del modello sarà quella di rendere esplicita la relazione di dipendenza fra queste due variabili rappresentate dai nodi H e E . Il problema essenziale è stabilire se la freccia che li congiunge e che rappresenta la dipendenza tra i due, si dirigerà dall'ipotesi H verso l'evidenza E (ammettendo dunque che la conoscenza dell'ipotesi H influenzi le probabilità associate agli stati della variabile evidenza E), oppure dall'evidenza E verso l'ipotesi H (ammettendo, viceversa, che la conoscenza dello stato della variabile E influenzi le probabilità associate agli stati dell'ipotesi H). È ragionevole assumere che la conoscenza dello stato della variabile H , produca determinate aspettative circa il possibile stato della variabile evidenza E . Se noi sapessimo che il sospetto è all'origine della traccia ritrovata, allora sarebbe naturale attendersi che il risultato dell'analisi del DNA metta in evidenza una corrispondenza tra i profili genetici.

Questo modello è rappresentato nella *Figura 1*.

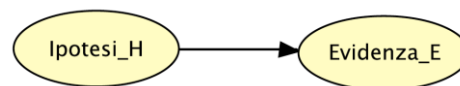


Figura 1: Rete Bayesiana composta da due nodi: le variabili H (ipotesi) e E (evidenza). La freccia che lega le due variabili (da H verso E) indica una dipendenza probabilistica tra le due variabili. La variabile H influenza la variabile E e viceversa.

Al fine di rendere operativa la rete per effettuare dei calcoli probabilistici, l'utilizzatore dovrà specificare le probabilità associate a tutti gli stati delle variabili.

Gli stati della variabile ipotesi H sono due e non essendovi frecce che influiscono sulla stessa, l'operatore dovrà specificare le probabilità (incondizionate) per gli stati 'vero' e 'falso' di questa variabile, $\Pr(H = \text{'vero'})$ e $\Pr(H = \text{'falso'})$. Immaginiamo, esclusivamente a scopo illustrativo, che la Corte sia in possesso di informazioni sul caso che rendano plausibile l'equiprobabilità dei due stati, vale a dire che sia ragionevole assumere $\Pr(H = \text{'vero'}) = \Pr(H = \text{'falso'}) = 0.5$.

Per quello che concerne i due stati della variabile E , le probabilità da specificare saranno quattro poiché ogni stato della variabile E dipende da ogni stato della variabile H (vi è una freccia che parte dal nodo H verso il nodo E). Sarà dunque necessario quantificare le seguenti quattro probabilità:

$\Pr(E = \text{'vero'} \mid H = \text{'vero'})$, $\Pr(E = \text{'falso'} \mid H = \text{'vero'})$, $\Pr(E = \text{'vero'} \mid H = \text{'falso'})$ e $\Pr(E = \text{'falso'} \mid H = \text{'falso'})$. Per chiarire il calcolo nel contesto del nostro esempio, consideriamo le seguenti possibili assegnazioni:

1. $\Pr(E = \text{'vero'} \mid H = \text{'vero'}) = 1$. Infatti, se il sospetto è all'origine della traccia ritrovata sul luogo del crimine ($H = \text{'vero'}$), allora è ragionevole supporre che, in assenza di errori nella tipizzazione del profilo genetico, il laboratorio dovrà forzatamente osservare una corrispondenza tra i profili genetici ($E = \text{'vero'}$).
2. $\Pr(E = \text{'falso'} \mid H = \text{'vero'}) = 0$. Seguendo il ragionamento fatto al punto 1., escludendo quindi un errore di laboratorio, l'evento che non vi sia corrispondenza genetica se il soggetto è all'origine della traccia è un evento impossibile. In termini matematici, osservando che i due eventi ($E = \text{'vero'}$) e



11/2018

($E = \text{'falso'}$) sono esaustivi, la probabilità cercata non è altro che il complemento a 1 della probabilità $\Pr(E = \text{'vero'} \mid H = \text{'vero'})$, vale a dire:

$$\Pr(E = \text{'falso'} \mid H = \text{'vero'}) = 1 - \Pr(E = \text{'vero'} \mid H = \text{'vero'}) = 1 - 1 = 0.$$

3. $\Pr(E = \text{'vero'} \mid H = \text{'falso'}) = 0.01$. La probabilità di ottenere una corrispondenza genetica tra i profili esaminati ($E = \text{'vero'}$) se il sospetto non è all'origine della traccia ritrovata sui luoghi del crimine ($H = \text{'falso'}$) è posta pari a 0.01. Da dove viene questo numero? Se il sospetto non è all'origine della traccia, allora un'altra persona, che possiede il medesimo profilo genetico del sospetto, sarà all'origine della traccia. Qual è la probabilità di osservare una persona che possieda questo profilo genetico? Ammettendo che sia plausibile considerare come popolazione di riferimento la popolazione italiana, l'operatore potrà fare ricorso ad un'opportuna collezione (banca dati) di profili genetici raccolti su vari campioni conosciuti al fine di assegnare un valore a questa probabilità. Supponiamo che, dopo aver consultato una banca dati relativa alla popolazione italiana, l'operatore osservi una frequenza relativa pari a 0.01 (l'1% dei profili presenti nella banca dati corrisponde a questo specifico profilo genetico). Questo valore potrà essere utilizzato per assegnare un valore alla probabilità cercata.

4. $\Pr(E = \text{'falso'} \mid H = \text{'falso'}) = 0.99$. Quest'ultima probabilità è stata ottenuta facendo un ragionamento analogo a quello condotto al punto 2. La probabilità cercata non è altro che il complemento a 1 della probabilità precedente, vale a dire:

$$\Pr(E = \text{'falso'} \mid H = \text{'falso'}) = 1 - \Pr(E = \text{'vero'} \mid H = \text{'falso'}) = 1 - 0.01 = 0.99.$$

Ora, alla Corte interessa conoscere la probabilità che il sospetto sia all'origine della traccia a partire dai risultati E emersi dalle analisi di laboratorio, vale a dire $\Pr(H = \text{'vero'} \mid E = \text{'vero'})$, mentre il laboratorio è in grado di elicitarne la probabilità di osservare una corrispondenza, ($E = \text{'vero'}$), nel caso in cui il soggetto sia o non all'origine della traccia, vale a dire $\Pr(E = \text{'vero'} \mid H = \text{'vero'})$ e $\Pr(E = \text{'vero'} \mid H = \text{'falso'})$. La necessaria inversione di probabilità può essere effettuata attraverso il Teorema di Bayes, che permette di aggiornare la probabilità, detta *a priori*, sugli stati della variabile ipotesi H alla luce dell'osservazione E fornita dal laboratorio.

Partendo da una probabilità che il sospetto sia all'origine della traccia pari a 0.5 (cioè 50%), come cambierà questo valore nel momento in cui si rende disponibile il risultato del laboratorio che mette in evidenza una corrispondenza genetica? Cioè, come possiamo passare dalla probabilità $\Pr(H = \text{'vero'})$ alla probabilità $\Pr(H = \text{'vero'} \mid E = \text{'vero'})$? Questo aggiornamento può essere effettuato attraverso il Teorema di Bayes nel modo seguente:

$$\Pr(H = \text{'vero'} \mid E = \text{'vero'}) = \Pr(H = \text{'vero'}) \times \Pr(E = \text{'vero'} \mid H = \text{'vero'}) / \Pr(E = \text{'vero'}),$$

dove

$$\Pr(E = \text{'vero'}) = \Pr(H = \text{'vero'}) \times \Pr(E = \text{'vero'} \mid H = \text{'vero'}) + \Pr(H = \text{'falso'}) \times \Pr(E = \text{'vero'} \mid H = \text{'falso'}).$$

Sostituendo i valori nella formula di Bayes otterremo:

$$\Pr(H = \text{'vero'} \mid E = \text{'vero'}) = 0.5 \times 1 / (0.5 \times 1 + 0.5 \times 0.01) = 0.9901.$$

La probabilità che lo stato dell'ipotesi H sia 'vero', cioè che il sospetto sia all'origine della traccia ritrovata sapendo che una corrispondenza genetica è stata osservata dal perito è ora pari a 0.9901, cioè circa il 99%. Il Teorema di Bayes costituisce un risultato fondamentale che consente di aggiornare la probabilità di interesse per la Corte in funzione delle evidenze raccolte e del loro valore nel contesto in esame. Da un valore della probabilità *a priori* (prima dell'acquisizione dell'evidenza) di 0.5, si passa ad una probabilità *a posteriori* (dopo l'acquisizione dell'evidenza) di 0.99. Se senza la prova scientifica credevamo che il sospetto fosse all'origine della traccia con una probabilità pari al 50%, ora, dopo l'acquisizione della prova scientifica, questa probabilità aumenta a 99%.

Vediamo ora come questa metodologia possa essere utilizzata per analizzare il caso in esame, vale a dire al fine di assegnare un valore alla probabilità che la collana sia stata strappata alla vittima dalla persona sospettata ($H = \text{'vero'}$) sapendo che un primo laboratorio (con una serie di metodiche analitiche particolari) ha osservato l'assenza di DNA e che un secondo laboratorio (con una serie di metodiche analitiche particolari, diverse dal primo) ha osservato la presenza di deboli quantità di DNA compatibile con quello della vittima.

5. L'analisi congiunta delle evidenze: misura del conflitto e soluzione.

Nel caso in esame, lo stato denominato H_p (ipotesi dell'accusa, $H = \text{'vero'}$) della variabile H corrisponde a 'La sospetta ha strappato la collana dal collo della vittima durante l'aggressione', mentre lo stato denominato H_d (ipotesi della difesa, $H = \text{'falso'}$) può essere enunciato come 'La collana appartiene alla sospetta, che non l'ha toccata per anni e non ne ricorda la provenienza'. Le evidenze, cioè i risultati prodotti dai due laboratori, sono rispettivamente 'Nessun profilo genetico è stato ottenuto dall'analisi della collana' e 'Una minima quantità di DNA compatibile con quello della vittima è stato messo in evidenza sulla collana'. Il problema iniziale per la costruzione del modello sarà quello di legare tra loro le diverse variabili in gioco. Il legame tra le ipotesi proposte dalle due parti e le evidenze risultanti dalle analisi dei due laboratori non è così semplice e diretto come invece illustrato nella *Figura 1* (una freccia dal nodo H verso il nodo E), sia perché vi sono molteplici evidenze, sia perché le ipotesi di interesse esprimono delle azioni – da un lato, quella di strappare, dall'altro, un'assenza di azione (la collana, secondo la difesa, non è stata toccata per anni) – e tali azioni causeranno delle conseguenze. Quali sono queste conseguenze?

La Rete Bayesiana da noi proposta per il caso in questione, e ora brevemente descritta¹³, è illustrata nella *Figura 2*. Vediamone gli elementi essenziali. Se una persona strappa una collana dal collo di un'altra persona, l'azione provocherà il trasferimento di materiale biologico dalla vittima alla collana, che si andrà ad aggiungere a quello già presente sulla collana a seguito del suo utilizzo, e di materiale biologico dalla persona che strappa la collana. Al contrario, se, come da ipotesi H_d , la collana non è utilizzata da anni, non ci si aspetta la presenza di materiale biologico trasferito durante l'azione e, soprattutto, non ci si aspetta la presenza di nessun, o poco, materiale genetico preesistente. Il modello proposto non può ignorare queste informazioni. I nodi denominati 'Transfer' e 'Background' (*Figura 2*) tengono conto, rispettivamente, del trasferimento di materiale relativo alla vittima ed al sospetto e della preesistenza di materiale biologico sulla collana.

Il DNA a disposizione del primo laboratorio per le analisi (DNA tempo t_1) dipenderà dalla quantità presente o trasferita al momento dello strappo (DNA tempo t_0), tenendo conto della persistenza di materiale genetico - che dipende principalmente dalla distanza temporale con l'azione che ha provocato il deposito del materiale genetico - al netto di una eventuale operazione di pulizia da parte della sospetta (Pulizia Collana). Va inoltre aggiunto che il test analitico effettuato dal primo laboratorio segue in verità un test precedente (Test indicativo), effettuato con l'obiettivo di individuare la presenza di cellule sulla collana. Questo test indicativo ha dato un risultato negativo. Il prelievo effettuato dal primo perito sulla collana per effettuare i test genetici provocherà ovviamente una diminuzione del materiale genetico potenzialmente ancora presente sulla collana dopo lo strappo e a disposizione del secondo perito (DNA tempo t_2).

Il modello proposto tiene inoltre in considerazione l'efficienza del metodo applicato dai due laboratori nonché della proporzione di superficie campionata ('Efficienza metodo' e 'Proporzione'). Il secondo laboratorio, quando effettuerà i prelievi sulla collana, per poter procedere alla seconda analisi di tipizzazione, applicherà un'altra metodologia con caratteristiche di efficacia proprie per utilizzare la quantità residua potenzialmente presente sulla collana.

Tutti questi fattori sono stati presi in considerazione, illustrati e quantificati nella Rete Bayesiana proposta (*Figura 2*).

Come presentato in precedenza con l'esempio illustrativo della *Figura 1*, ogni nodo (variabile) è composto da due o più stati. Il nostro modello presenta anche variabili con più di due stati perché molti nodi rappresentano quantità di materiale biologico raccolto o osservato. Per esempio, il DNA presente sulla collana al momento del prelievo fatto dal primo perito per l'analisi indicativa presenta 10 stati che descrivono la quantità di materiale genetico (assenza; picogrammi, nanogrammi o quantità più importanti di DNA provenienti, rispettivamente, dalla vittima, dalla persona sospettata o da una mistura delle due persone). Il nodo 'Test indicativo', che avrà gli stati 'Risultato positivo (cioè presenza di DNA)' o 'Risultato negativo (assenza di DNA)' sarà condizionato dagli

¹³ Una descrizione dettagliata del modello e dei valori utilizzati per quantificare le tavole di probabilità condizionate è presentata nel testo I. DE MARCH, F. TARONI, *Bayesian networks for dissonant evidence: a case study*, Internal Report, Ecole des sciences criminelles, Université de Lausanne, 2018.

stati della variabile 'DNA a t₁' (cioè la quantità di DNA appena prima del prelievo). La probabilità che il risultato dell'analisi indicativa sia positivo se lo stato 'Assenza di DNA' è vero viene posta pari a 0. La probabilità che il risultato dell'analisi indicativa sia positivo se lo stato 'Presenza di picogrammi di DNA della vittima' è vero, è posta pari a 0,8, infine, la probabilità che il risultato dell'analisi indicativa sia positivo se lo stato 'Presenza di nanogrammi di DNA della vittima' è vero, cioè una quantità più importante di DNA della vittima è presente è posta pari a 1 (evento certo). Le tabelle di probabilità condizionate presenti nella Rete Bayesiana che rappresenta il nostro modello, sono state tutte compilate facendo riferimento alle informazioni e ai dati numerici disponibili a partire da ricerche scientifiche pubblicate o test di laboratorio.

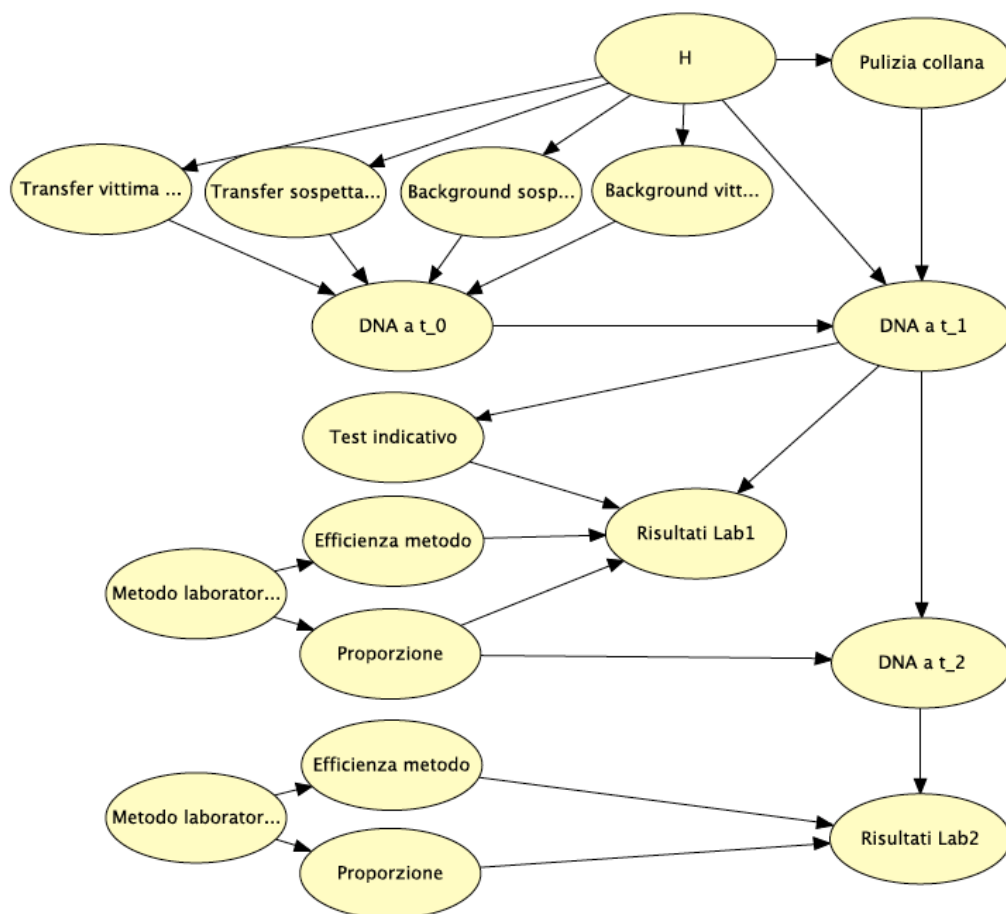


Figura 2: Rete Bayesiana che descrive il caso d'interesse dove un test indicativo sulla quantità di DNA presente sulla collana e due risultati genetici di tipizzazione sono presi in considerazione ('Test indicativo', 'Risultati Lab1' e 'Risultati Lab2', rispettivamente). Tutte le variabili di interesse incluse nel modello proposto sono descritte nel testo.



11/2018

Quale sarà dunque l'apporto informativo di questo modello, una volta che saranno inserite le evidenze disponibili, cioè (1) il risultato negativo del test indicativo, (2) il risultato dell'assenza di DNA sulla collana osservato dal primo laboratorio e (3) il risultato positivo di presenza di DNA compatibile con quello della vittima espresso dal secondo laboratorio?

Il modello da noi proposto permette di calcolare la probabilità associata alle ipotesi fondamentali d'interesse. Ma prima di soffermarci sul risultato ottenuto, dobbiamo sottolineare che il modello permette ugualmente di allertare l'utilizzatore qualora appaiano problemi specifici idonei a rendere problematica l'attività inferenziale. Nel caso specifico, il fatto di avere due risultati dissonanti, dove il primo sostiene uno stato particolare dell'ipotesi, e il secondo sostiene lo stato alternativo della stessa, crea una tensione nel modello che si esprime con un valore, detto di 'conflitto'¹⁴. Il valore ottenuto dal nostro modello indica la presenza di un conflitto, non chiarito, tra i risultati osservati. Il fatto che i due laboratori utilizzino metodiche differenti non annulla questo valore di conflitto. Il modello proposto tiene infatti in debito conto le differenze fra le metodiche dei due laboratori (nodi 'Efficienza metodo' e 'Proporzione' associati alle metodiche dei due laboratori).

L'esistenza o il riconoscimento di un conflitto tra i risultati è di fondamentale importanza, poiché esso è in grado di minare, se non risolto, l'attendibilità del risultato finale (vale a dire le probabilità a posteriori delle ipotesi di interesse per la Corte). Un modo per risolvere il conflitto originato da risultati contrastanti forniti dai laboratori di genetica forense consiste nell'aggiungere una nuova variabile al nostro modello (si veda la Figura 3, variabile 'Contaminazione'). Si tratta di una variabile che fa riferimento alla possibile contaminazione del reperto con materiale genetico proveniente dalla vittima (come illustrato nel nostro primo articolo), e di seguito indicata con la lettera V . Gli stati della variabile sono i seguenti: nessuna contaminazione ('0'), una contaminazione di picogrammi (pg) di sangue proveniente dalla vittima (V, pg), una contaminazione di nanogrammi (ng) di sangue proveniente dalla vittima (V, ng), una contaminazione abbondante da sangue proveniente dalla vittima ($V, >ng$). Le probabilità associate a questi stati della variabile sono le seguenti: $Pr(0) = 0.99$; $Pr(V, pg) = 0.01$; $Pr(V, ng) = 0$; $Pr(V, >ng) = 0$.

Questi valori sottolineano la forte credenza che non vi siano state contaminazioni. In effetti, alla probabilità relativa al primo stato della variabile 'Contaminazione', quello che stipula l'impossibilità di una contaminazione, è assegnato un valore pari a 0.99. Da questo, segue che la restante probabilità di 0.01, dovrà essere redistribuita tra gli stati che ammettono una qualche quantità di materiale genetico proveniente dalla vittima. In realtà, una probabilità nulla viene assegnata agli stati che indicano una contaminazione maggiore, e lo 0.01 viene interamente attribuito allo stato che accoglie una contaminazione con una piccolissima quantità (picogrammi, pg) di

¹⁴ F.V. JENSEN, *Bayesian Networks and Decision Graphs*, Springer, New York 2001; U.B. KJÆRULFF, A.L. MADSEN, *Bayesian networks and Influence Diagrams – A Guide to Construction and Analysis*, Springer, New York 2008.

DNA proveniente dalla vittima. Si accoglie pertanto, favorendo la visione accusatoria, la tesi che la contaminazione sia poco probabile.

L’inserimento di questa nuova variabile consente di risolvere il conflitto. Anziché escludere il risultato del primo laboratorio, in conflitto con il secondo, si procede considerando tutte le informazioni disponibili (i tre risultati, due dal primo laboratorio ed uno dal secondo), inclusa la reale possibilità, anche se remota, di una avvenuta contaminazione dei reperti. Solo a questo punto, è possibile calcolare le probabilità per le variabili di interesse.

In funzione dei valori inseriti nelle tavole di probabilità, basati su dati provenienti da banche dati o pubblicati nella letteratura scientifica, la probabilità che la collana sia stata strappata dalla persona sospettata, alla luce dei tre risultati di analisi è nettamente inferiore alla probabilità che la collana sia sempre appartenuta alla sospetta (20% di probabilità contro 80% di probabilità che la collana appartenga alla persona sospettata).

I risultati delle analisi dei due laboratori, se considerati congiuntamente, supportano la tesi della Difesa e questo in contraddizione con quanto affermato dalle Corti che hanno dato unicamente credito al risultato del secondo laboratorio escludendo categoricamente le osservazioni fatte dal primo laboratorio.

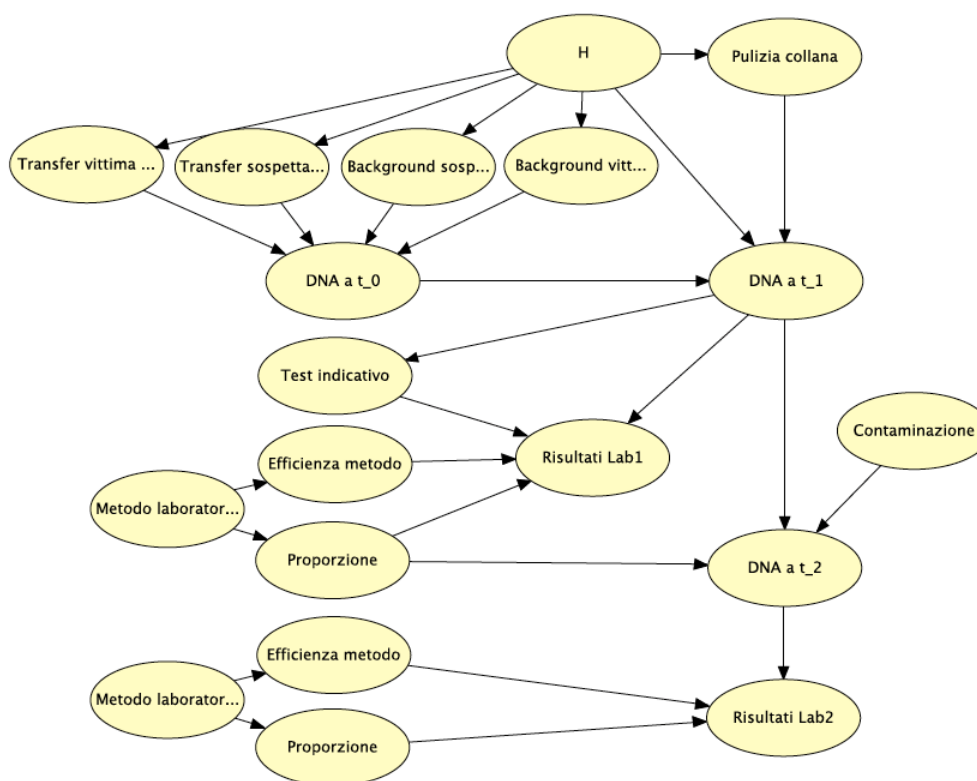


Figura 3: Rete Bayesiana che descrive il caso d’interesse dove un test indicativo sulla quantità di DNA presente sulla collana e due risultati genetici di tipizzazione sono presi in considerazione (‘Test indicativo’, ‘Risultati Lab1’ e ‘Risultati Lab2’, rispettivamente). La Rete è modificata rispetto a quella presentata nella Figura 2 per includere il nodo ‘Contaminazione’, introdotto per spiegare la presenza di conflitto fra i diversi risultati.

6. Conclusione.

Il caso esaminato e discusso dai giudici di due Corti di Assise e da quelli dalla Suprema Corte di Cassazione mette in evidenza – a parere di chi scrive – una gestione non ottimale della problematica della valutazione congiunta di elementi di prova scientifici, che nel caso d’interesse sono i risultati dissonanti emersi da analisi genetiche condotte da due laboratori.

Come gestire congiuntamente l’informazione prodotta da un primo laboratorio che, dopo gli esiti negativi sulla presenza di tracce biologiche sul reperto, ne mette in evidenza altri, ugualmente negativi, sulla tipizzazione dei profili genetici, e quella fornita da un secondo laboratorio che riporta la presenza di una debole quantità di DNA corrispondente al profilo genetico della vittima?

I giudici hanno sostenuto che i risultati del primo laboratorio non debbano essere presi in considerazione: le metodiche di estrazione del DNA impiegate dal secondo laboratorio sono state giudicate più efficaci e dunque il risultato ottenuto dal secondo laboratorio è stato considerato l’unico attendibile. Quanto riportato dal primo laboratorio è stato pertanto escluso, in virtù della supposta inadeguatezza della metodologia adottata ai fini dell’inchiesta.

Noi siamo dell’opinione che tutti i risultati disponibili debbano essere interpretati congiuntamente, tenendo ovviamente in debito conto delle metodiche di estrazione e di tipizzazione eseguite dai vari laboratori (e dunque della loro maggiore o minore efficacia). Ai fini di una corretta valutazione, vi sono altre variabili d’interesse che devono ugualmente essere prese in considerazione, tipicamente i fenomeni di trasferimento e la persistenza di materiale biologico durante le azioni contestate e i prelievi utili per le analisi di laboratorio. In effetti, escludere un risultato sulla base della presunta inefficacia della metodologia implementata dal laboratorio, priverà il decisore di informazioni fondamentali per il verdetto; questo è quanto abbiamo cercato di illustrare in questo articolo.

Cosa ci mostra l’analisi congiunta dei risultati genetici? L’analisi probabilistica quantifica l’evidente conflitto esistente fra i diversi risultati. Il problema da affrontare, prima di utilizzare un eventuale modello per quantificare la probabilità dell’ipotesi d’interesse e in seguito prendere una decisione giuridica, sarà quello di risolvere il conflitto messo in evidenza dai risultati delle analisi di laboratorio.

La Difesa¹⁵, in effetti, aveva in precedenza messo in evidenza una serie di osservazioni volte a sostenere l’idea che non fosse affatto possibile escludere una contaminazione della collana con del sangue proveniente da reperti prelevati sul luogo del crimine, offrendo una spiegazione alternativa e logica alla presenza del profilo genetico della vittima sulla collana (presenza che non è mai stata contestata dalla Difesa).

¹⁵ F. TARONI, S. BOZZA, P. GARBOLINO, *Contaminazioni di un reperto con il DNA*, cit., mette in evidenza la lista di aspetti critici che non permettono di escludere che una contaminazione sia effettivamente avvenuta.



11/2018

Questo punto specifico non è stato accolto dalle diverse Corti, ed i dubbi sollevati dalla Difesa in sede dibattimentale sui fattori che avrebbero potuto originare una contaminazione, non hanno trovato risposta.

La variabile 'Contaminazione', se introdotta nel modello illustrato nella figura 2, permette (anche se quantificata probabilisticamente in favore dell'Accusa)¹⁶ di risolvere il conflitto e dunque di quantificare in modo congruo le probabilità associate alle tesi proposte, rispettivamente, dall'Accusa e dalla Difesa. Il risultato ottenuto, tenendo conto dei risultati dissonanti e delle diverse metodiche di laboratorio (ciascuna con un diverso grado di affidabilità), favorisce la tesi della Difesa. La figura 3 illustra il modello finale. Questo modello permette dunque di valutare congiuntamente le informazioni genetiche disponibili e possiede inoltre il vantaggio di essere flessibile e evolutivo e poter dunque accogliere e modellizzare, se necessario, altre informazioni raccolte durante l'inchiesta.

Cosa significa aver osservato delle tracce di suole di scarpe nel sangue all'interno dell'appartamento della vittima, le cui caratteristiche non corrispondono a quelle delle scarpe in possesso della persona sospettata? Cosa significa non aver osservato nessuna traccia di contatto (tracce digitali, DNA, fibre tessili, ecc.) della persona sospettata all'interno dell'appartamento della vittima e sulla vittima medesima?

La crescente complessità dell'analisi prodotta dall'aumento delle informazioni disponibili, può essere gestita con successo attraverso lo sviluppo della Rete Bayesiana proposta e la conseguente introduzione di nuove variabili. Il grande vantaggio di questo approccio consiste nel limitare la gestione della complessità alla soluzione dell'innegabile necessità di mettere in evidenza, operazione non sempre facile, i legami di dipendenza esistenti tra le variabili (e dunque, la presenza di frecce tra i diversi nodi e le tabelle di probabilità a loro collate); le probabilità di interesse saranno invece automaticamente aggiornate alla luce delle informazioni inserite.

¹⁶ Si ricorda che una probabilità dell'1% è stata attribuita allo stato che esprime la potenziale contaminazione con deboli quantità di materiale biologico. È stata espressa l'impossibilità che una contaminazione con quantità maggiori di materiale biologico sia avvenuta e una probabilità del 99% è stata assegnata allo stato della variabile che recita 'Assenza di contaminazione'.