AI Sherlock Holmes Style

Introduction to Automated Abductive Inference

•••

Phillip Rhodes
All Things Open
Raleigh NC
10-15-2019

“Abductive” Inference?

As opposed to what?

Inference?

- Inferences are steps in reasoning, moving from premises to logical consequences
- In epistemology, a number of specific methods for acquiring new knowledge are defined: perception, introspection, memory, **reason**, and testimony
- The distinction between “reasoning” and “inference” is subtle and we’ll mostly treat the terms as synonymous

“That’s a strange thing, you are the second man today that has used that expression to me.” -- Stamford

“And who was the first?” -- Watson

“A fellow who is working at the chemical laboratory at the hospital … He is a little queer in his ideas - an enthusiast in some branches of science” -- Stamford

“A medical student, I suppose?” -- Watson

(A Study in Scarlet)

Reasoning

- Reasoning is a means to:
  - Establish and verify facts
  - Make sense of things
  - Adapt or justify practices and beliefs
  - Explain or justify events
  - Predict the future
  - Create new knowledge
  - Build a model of the world from sensory perception

Types of Reasoning / Inference

- Classical Logic
- Statistical
    - Bayesian
    - Frequentist
- Analogical
- Spatial
- Temporal
- Prediction

“You see Watson, it was perfectly obvious from the first that the only possible object of this rather fantastic business of the advertisement of the League, and the copying of the encyclopedia, must be to get this not over-bright pawnbroker out of the way for a number of hours every day.” -- Sherlock Holmes

(The Red Headed League)

Classical Logic

- Traditionally (since ~300 BC) divided into two main branches:
  - Deductive
  - Inductive

Deductive Inference

- If a Deductive argument is **valid** and the premises are true, the conclusion **must** be true
- Syllogism

All men are mortal

Socrates is a man

Socrates is mortal

“You know me too well to think that I am boasting when I say that I shall either confirm or destroy his theory by means which he is quite incapable of employing, or even understanding. To take the first example at hand, I very clearly perceive that in your bedroom the window is on the right-hand side, and yet I question whether Mr. Lestrade would have noted even so self-evident a thing as that. -- Holmes

“How on Earth??” -- Watson

“My dear fellow, I know you well. I know the military neatness which characterizes you. You shave every morning, and in this season you shave by the sunlight; but since your shaving is less and less complete as we get farther back on the left side, until it becomes positively slovenly as we get round the angle of the jaw, it is surely very clear that that side is less illuminated than the other.” -- Holmes

(The Boscombe Valley Mystery)

More Deduction

- Modus Ponens
- X □ Y, X : Y

X □ Y

X

Y

More Deduction

- Modus Tollens
- X □ Y, ~Y : ~X

X □ Y

~Y

~X

“You appeared to be surprised when I told you, on our first meeting, that you had come from Afghanistan.” - Sherlock Holmes

“You were told no doubt.” -- Watson

“Nothing of the sort. I *knew* you had come from Afghanistan... “The train of reasoning ran: ‘Here is a gentleman of a medical type, but with the air of a military man. Clearly an Army doctor then. He has just come from the tropics for his face is dark, and that is not the natural tint of his skin, for his wrists are fair. He has undergone hardship and sickness, as his haggard face says clearly. His left arm has been injured.He holds it in a stiff and unnatural manner. Where in the tropics could an English army doctor have seen much hardship and got his arm wounded? Clearly in Afghanistan.’”

(A Study in Scarlet)

Deduction Can Get Complicated

- Remember these crazy truth tables from Discrete Math?

P Q	P	□	(~	Q	□	(P	&	~	Q))
T T	T	T	F	T	T	T	F	F	T
T F	T	T	T	F	T	T	T	T	F
F T	F	T	F	T	T	F	F	F	T
F F	F	T	T	F	F	F	F	T	F

Inductive Inference (Generalisation)

- Inductive Generalisation
- Infer from some set of observations, some general truth about the observed thing

This person in this town is nice

This person in this town is nice

This person in this town is nice

This person in this town is nice

Everybody in this town is nice

“I have already explained to you that what is out of the common is usually a guide rather than a hindrance. In solving a problem of this sort, the grand thing is to be able to reason backward. That is a very useful accomplishment, and a very easy one, but people do not practice it much. In the everyday affairs of life it is more useful to reason forward, and so the other comes to be neglected. There are fifty who can reason synthetically for one who can reason analytically.” - Sherlock Holmes

(A Study in Scarlet)

“The Problem of Induction”

- A single counter example completely destroys your inductive argument, no matter how many positive examples you have
- If the argument is couched in terms of “All X”
- “All swans are white”
- Was accepted scientific wisdom... until a species of black swans were discovered in New Zealand
- A notorious problem that lies at the heart of all inductive reasoning
- Major thorn in the side of empiricists everywhere
- This is why induction is generally considered a “weaker” form of inference, compared to deduction

Examples of Statistical Reasoning

- Regression
- Using linear regression to fit a line to data
- y=mx+b
- Statistical Syllogism
- Can think of these as the probabilistic version of inductive generalization
- Rules couched in terms of “MOST X are …” instead of “ALL X are …”

“Most people, if you describe a train of events to them, will tell you what the result would be. They can put those events together in their minds, and argue from that something will come to pass. There are few people, however, who, if you told them a result, would be able to evolve from their own inner consciousness what the steps were which led up to that result. This power is what I mean when I talk of reasoning backwards, or analytically” -- Sherlock Holmes

(A Study in Scarlet)

Statistical Syllogism Continued

Almost all people are more than 2 feet tall

Rodney is a person

Rodney is more than 2 feet tall

Adding “Abductive” Inference To The Mix

- Charles Sanders Peirce began to define **abduction** as a unique form of inference - apart from induction - in the 1860’s.
- “Inference to the best explanation”
- A form of inference that goes from data describing something, to a hypothesis that best explains or accounts for the data.
- Peirce contended that abduction, as a distinctive pattern of reasoning wherein explanatory hypotheses are formed and accepted, occurs in science and in everyday life
- Charniak and McDermott characterize abduction as modus ponens turned backwards, inferring the cause of something; as generation of explanations for what we see around us

“To begin at the beginning, I approached the house, as you know, on foot, and with my mind entirely free from all impressions. I naturally began by examining the roadway, and there, as I have already explained to you, I saw clearly the marks of a cab, which, I ascertained by query, must have been there during the night. I satisfied myself that it was a cab and not a private carriage by the narrow gauge of the wheels. The ordinary London growler is considerably less wide than a gentleman’s brougham.” -- Sherlock Holmes

(A Study in Scarlet)

Abduction in Human Thought

- Medical diagnosis
- Criminal Investigation
- Fault diagnosis (machines, computer programs, etc.)
- Story understanding
- Vision
- Natural Language Understanding

The Abduction Pattern

D is a collection of data (facts, observations, givens)

H explains D (would, if true, explain D)

No other hypothesis can explain D as well as H does

____

Therefore, H is probably true

“Some facts should be suppressed, or, at least, a just sense of proportion should be observed in treating them. The only point in the case which deserved mention was the curious analytical reasoning from effects to causes, by which I succeeded in unravelling it” - Sherlock Holmes

(The Sign of Four)

Abduction in Ordinary Life

Joe: Why are you pulling into the filling station?

Tidmarsh: Because the gas tank is nearly empty.

Joe: What makes you think so?

Tidmarsh: Because the gas gauge indicates nearly empty. Also, I have no reason to think that the gauge is broken, and it has been a long time since I filled the tank.

"Nearly empty gas tank" is the best explanation for the gauge reading. An alternative hypothesis is that the gauge is broken, but there's no particular reason to believe this. OTOH, "long time since tank was filled" provides additional support for the "tank is nearly empty" hypothesis.

Abduction in Detection

Harman (1965) says that when a detective puts the evidence together and decides that the culprit must have been the butler, the detective is really reasoning that no other explanation that accounts for all the facts is plausible enough or simple enough to be accepted.

Truzzi (1983) alleges that at least 217 examples of abduction can be found in the Sherlock Holmes canon.

“There is no great mystery in this matter," he said, taking the cup of tea which I had poured out for him; “the facts appear to admit of only one explanation.” -- Sherlock Holmes

(The Sign of Four)

Abduction and Artificial Intelligence

- Deduction (various inference engines) and induction (via rule induction algorithms like C4.5, etc.) are already widely implemented on digital computers
- Implementing abductive inference on a computer gives us one more tool for building artificially intelligent systems
- First attempts to introduce abductive reasoning into AI began about 1973 (Pople)
- Early successes:
  - RED (antibody identification in blood)
  - RED-2
  - MDX2
  - TIPS
  - QUAWDS

“Ah, my dear Watson, there we come to the realms of conjecture where the most logical mind may be at fault. Each may form his own hypothesis upon the present evidence, and yours is as likely to be correct as mine.” -- Sherlock Holmes,

(The Empty House)

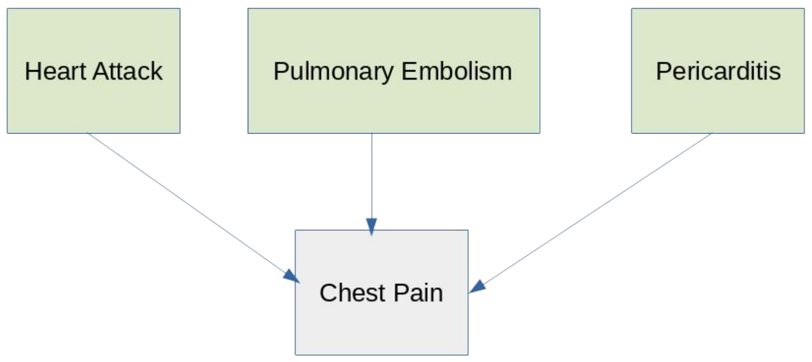
Implementing Abduction On A Computer

- Parsimonious Covering Theory (PCT)
- Major challenge: dealing with the combinatorial explosion of possible hypotheses (since effects may have multiple causes)
- Prescreening “cause fragments” to remove the implausible ones helps deal with this
- Evaluating the plausibility of different competing hypotheses is difficult
- “Parsimony criteria” are one way to disambiguate among possible hypotheses

Implementing Abduction, Continued

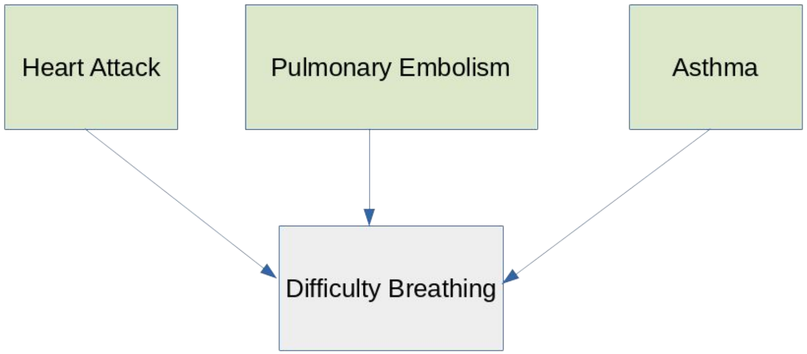
- Human abductive reasoning seems to involve stages of "hypothesis generation", "hypothesis updating", and "hypothesis testing", combined in a "hypothesize-and-test" cycle
- In PCT terms, the process involves three steps:
  - Disorder evocation - a set of individual hypothesis elements are evoked through association to a given manifestation
  - Hypothesis formation - the hypothesis elements are combined with previously formed hypotheses to form a set of new hypotheses that can account for both old and new manifestations
  - Question generation (or "hypothesis testing") - a new question is generated, whose answer may be used to test and further disambiguate existing hypotheses

Single Manifestation



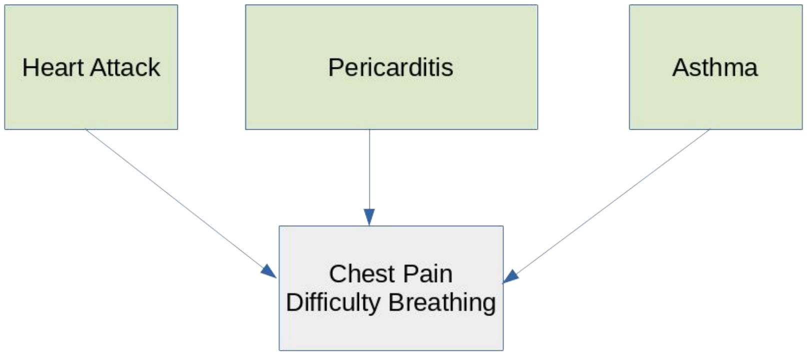
Hypothesis: “Heart Attack” OR “Pulmonary Embolism” OR “Pericarditis”

Single Manifestation 2



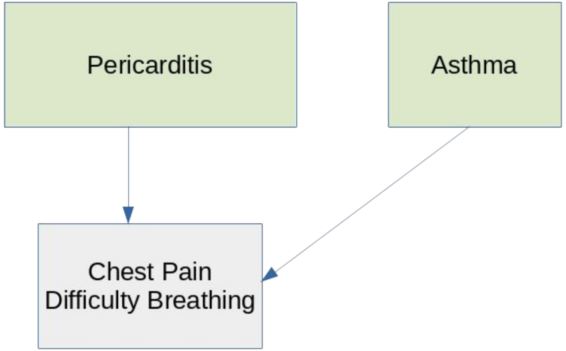
Hypothesis: “Heart Attack” OR “Pulmonary Embolism” OR “Asthma”

Multiple Manifestations
Redundant Hypothesis



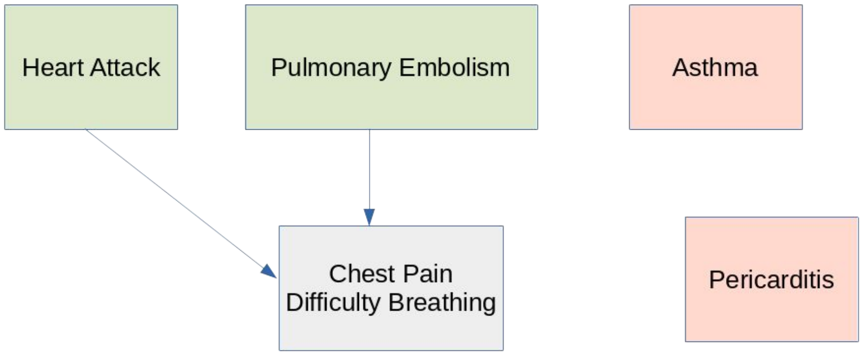
Hypothesis: “Heart Attack” AND “Pericarditis” AND “Asthma”

Multiple Manifestations
Non-Parsimonious Hypothesis



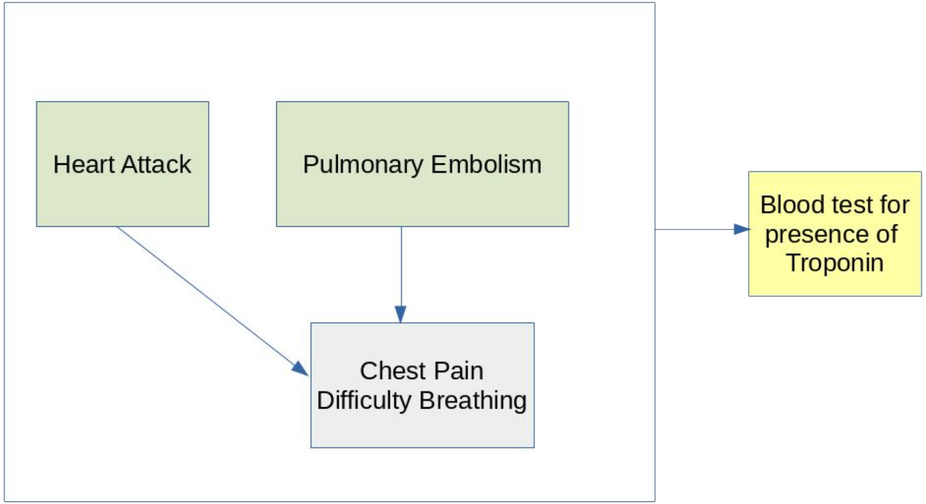
Hypothesis: “Pericarditis” AND “Asthma”

Multiple Manifestations
Two Parsimonous (But Competing) Hypotheses



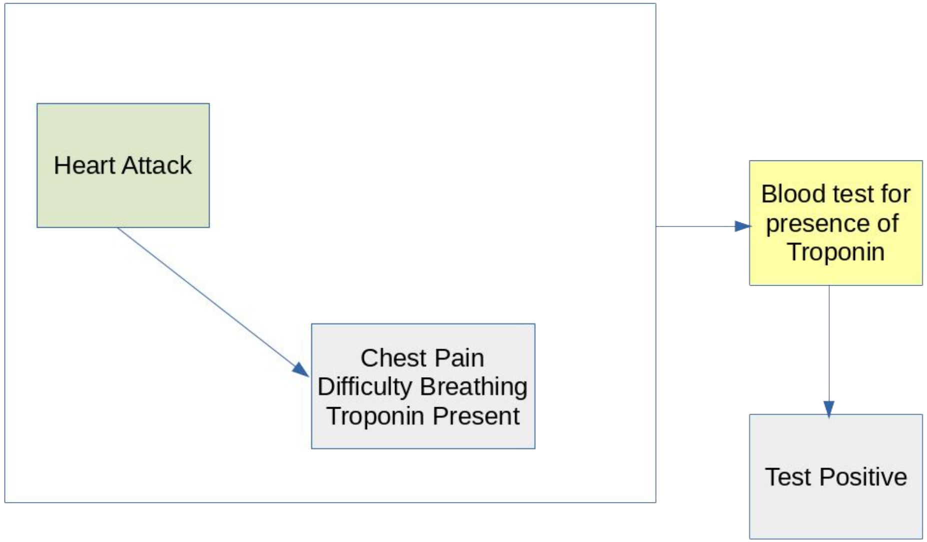
Hypothesis: “Heart Attack” OR “Pulmonary Embolism”

Needs Additional Information



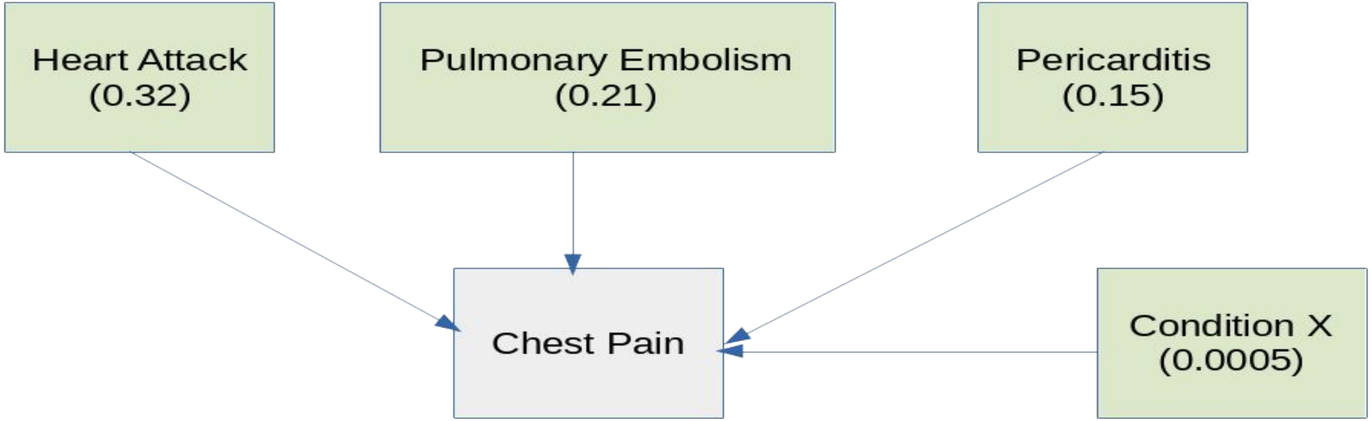
Question: Are enzyme markers showing cardiac tissue damage present?

Update Hypothesis



Hypothesis: Heart Attack

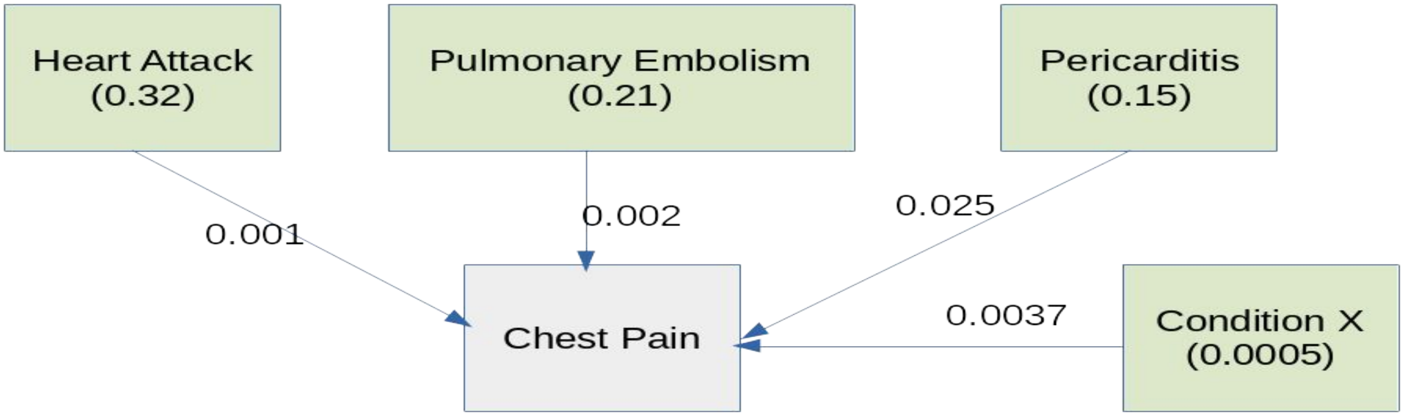
Introducing Probability



Hypothesis: “Heart Attack” OR “Pulmonary Embolism” OR “Pericarditis”

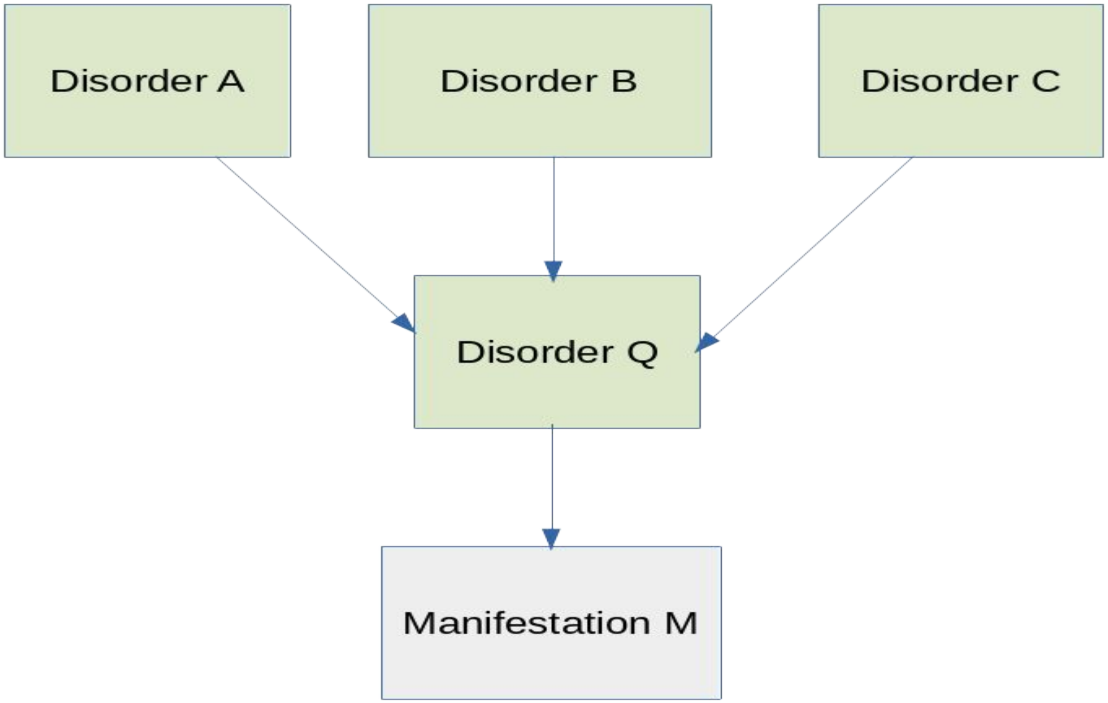
(we ignore Condition X because the prevalence is so low)

Introducing Probability 2



Given prevalence and priors, we can apply Bayes Theorem here

Causal Chains



Unfortunately...

A truly deep dive into the theoretical foundations of PCT would take far longer than the time we have available.

$causes(m_3) = \{d_2,d_3,d_5,d_6\}$ $\quad$ $causes(m_2) = \{d_5,d_6,d_7,d_9\}$

$causes(m_5) = \{d_7,d_8,d_9\}$ $\quad$ $causes(m_4) = \{d_1,d_2,d_8\}$

$causes(m_6) = \{d_2,d_4,d_8\}$

$G = \{G_1, G_2\}$ is a generator–set representing the set of all irredundant covers of $M^+ = \{m_1,m_4,m_5\}$ where $G_1 = (\{d_3,d_4\}\ \{d_8\})$ and $G_2 = (\{d_1,d_2\}\ \{d_7,d_8,d_9\})$

$H_1 = causes(m_3) = \{d_2, d_3, d_5, d_6\}$

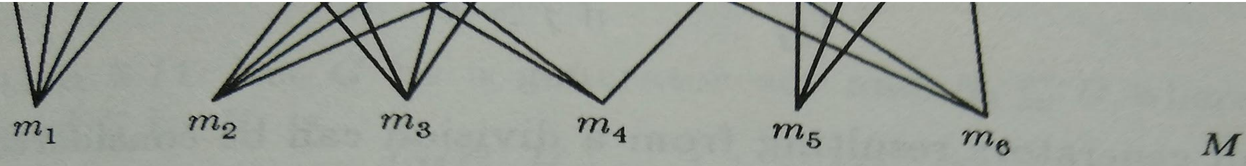
$Q^J = div(G, H_1) = Q^{J_1} \cup Q^{J_2}$ represents all irredundant covers of $\{m_1,m_4,m_5\}$ that also cover m_3 where $Q^{J_1} = div(G_1, H_1) = \{(\{d_3\}\ \{d_8\})\}$ and $Q^{J_2} = div(G_2, H_1) = \{(\{d_2\}\ \{d_7,d_8,d_9\})\}$

$Q^K = res(G, H_1) = Q^{K_1} \cup Q^{K_2}$ represents all irredundant covers of $\{m_1,m_4,m_5\}$ that do not cover m_3 where $Q^{K_1} = res(G_1, H_1) = \{(\{d_4\}\ \{d_8\})\}$ and $Q^{K_2} = res(G_2, H_1) = \{(\{d_1\}\ \{d_7,d_8,d_9\})\}$

$Q^L = augres(G, H_1) = Q^{L_1} \cup Q^{L_2}$ where $Q^{L_1} = augres(G_1, H_1) = \{(\{d_4\}\ \{d_8\}\ \{d_2,d_5,d_6\})\}$ and $Q^{L_2} = augres(G_2, H_1) = \{(\{d_1\}\ \{d_7,d_8,d_9\}\ \{d_3,d_5,d_6\})\}$

$Q^R = res(Q^L, Q^J) = \{(\{d_4\}\ \{d_8\}\ \{d_5,d_6\})(\{d_1\}\ \{d_7,d_8,d_9\}\ \{d_5,d_6\})\ (\{d_1$ represents all irredundant covers of $\{m_1,m_4,m_5,m_3\}$ that are redundant covers of $\{m_1,m_4,m_5\}$

$G^{rev} = Q^J \cup Q^R$ is a generator–set representing the set of



$causes(m_1) = \{d_1, d_2, d_3, d_4\}$ $causes(m_2) = \{d_5, d_6, d_7, d_9\}$
$causes(m_3) = \{d_2, d_3, d_5, d_6\}$ $causes(m_4) = \{d_1, d_2, d_8\}$
$causes(m_5) = \{d_7, d_8, d_9\}$ $causes(m_6) = \{d_2, d_4, d_8\}$

$G = \{G_1, G_2\}$ is a generator–set representing the set of all irredundant covers of $M^+ = \{m_1, m_4, m_5\}$
where $G_1 = (\{d_3, d_4\}\ \{d_8\})$ and $G_2 = (\{d_1, d_2\}\ \{d_7, d_8, d_9\})$

$H_1 = causes(m_3) = \{d_2, d_3, d_5, d_6\}$

$Q^J = div(G, H_1) = Q^{J_1} \cup Q^{J_2}$ represents all irredundant covers of $\{m_1, m_4, m_5\}$ that also cover m_3
where $Q^{J_1} = div(G_1, H_1) = \{(\{d_3\}\ \{d_8\})\}$
and $Q^{J_2} = div(G_2, H_1) = \{(\{d_2\}\ \{d_7, d_8, d_9\})\}$

$Q^K = res(G, H_1) = Q^{K_1} \cup Q^{K_2}$ represents all irredundant covers of

References

- *Abductive Inference Models for Diagnostic Problem-Solving* - Yun Peng & James Reggia, Springer-Verlag, 1990
- *Abductive Inference: Computation, Philosophy, Technology* - John R. Josephson & Susan G. Josephson, Cambridge University Press, 1996
- *The Complete Sherlock Holmes, Volume 1* - Sir Arthur Conan Doyle, Barnes & Noble Books, 2004
- *Logic for Dummies* - Mark Zegarelli, Wiley Publishing, 2007
- https://en.wikipedia.org/wiki/Abductive_reasoning
- https://en.wikipedia.org/wiki/Reason
- https://en.wikipedia.org/wiki/Inference
- https://en.wikipedia.org/wiki/Deductive_reasoning
- https://en.wikipedia.org/wiki/Inductive_reasoning
- https://en.wikipedia.org/wiki/Statistical_inference
- https://plato.stanford.edu/entries/epistemology/