On Conditional Probability Problem Solving Research — Structures and Contexts

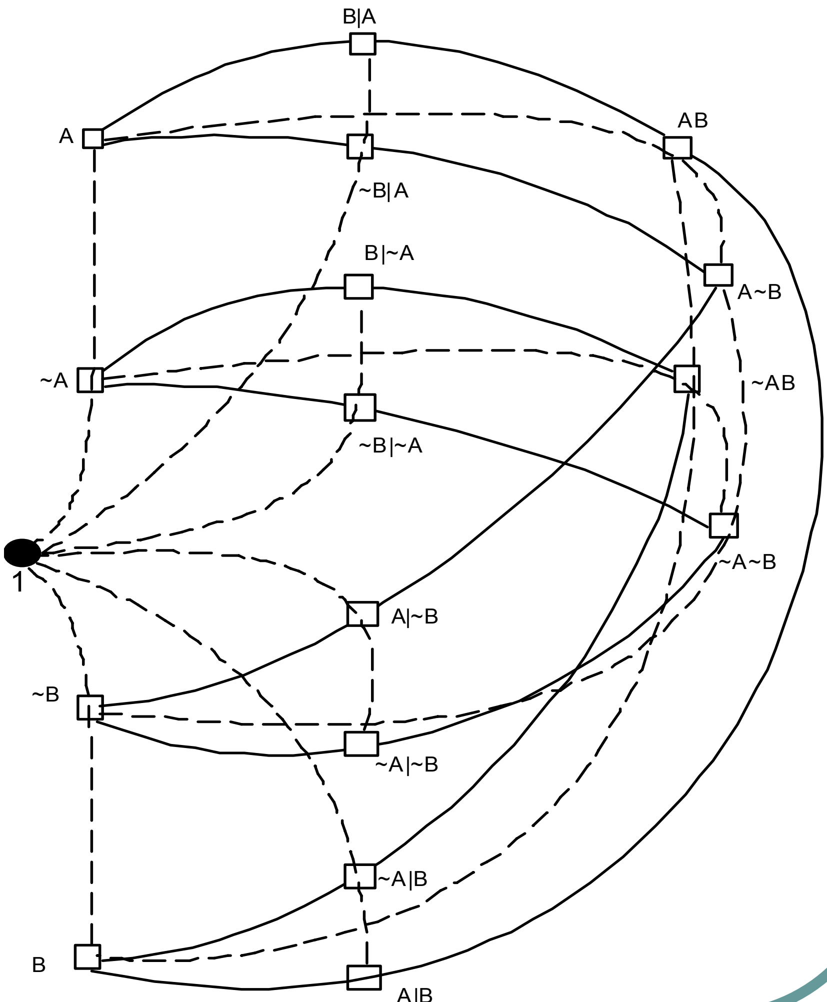
M. Pedro Huerta
Universitat de València
Spain

WHAT PROBLEMS? WHAT STRUCTURES?

- We investigate a particular world of (school and word) conditional probability problems.
- This world of problems is generated by
  - the events A, B, ~A, ~B, AB, A~B, ~A B, ~A~B;
  - 16 probabilities (4 absolute probabilities, 8 conditional probabilities and 4 intersection probabilities); and
  - 18 relationships between these probabilities (8 multiplicative relationships and 10 additive relationships).
- Basically, solving a problem (from this world) is a process to find an unknown probability when only three known and no directly-related probabilities are given.
- Consequently, posing a problem would consist in to formulate events, quantities (probabilities or not) and relationships (between events and quantities) in context, either mathematical or non-mathematical context.
- Ternary Problems of Conditional Probability can be initially described by means of a three-components vector (x, y, z) with x +y +z=3 and x, y, and z conveniently chosen.

WHAT TOOL AND METHOD FOR ANALYSING PROBLEMS?

- A mathematical object, the TRINOMIAL GRAPH (representing data and relationships between data)
- A problem solving method: The Analysis-Synthesis Method



PROBLEMS IN CONTEXT: SITUATIONS AND CONTEXTS

- A problem is an instance of something more general we call Situation.
- A Context is a particular situation in which problems can be put forward.
- Diagnostic Test is a situation and Diagnostic Test in Health is a context.
- In the Diagnostic Test Situation risks are conditional probabilities, but they have different meanings depending on the context where risks are assumed and with different consequences.
- Two well-known problems, *the disease problem* and the *taxi cab problem*, can be labelled as ternary problems of conditional probability, in the same situation (Diagnostic test) but in different context (Health, Assurance).

PROBLEMS IN CONTEXT: PHENOMENOLOGICAL ANALYSIS METHOD

In addition to trinomial graphs and the analysis-synthesis method,

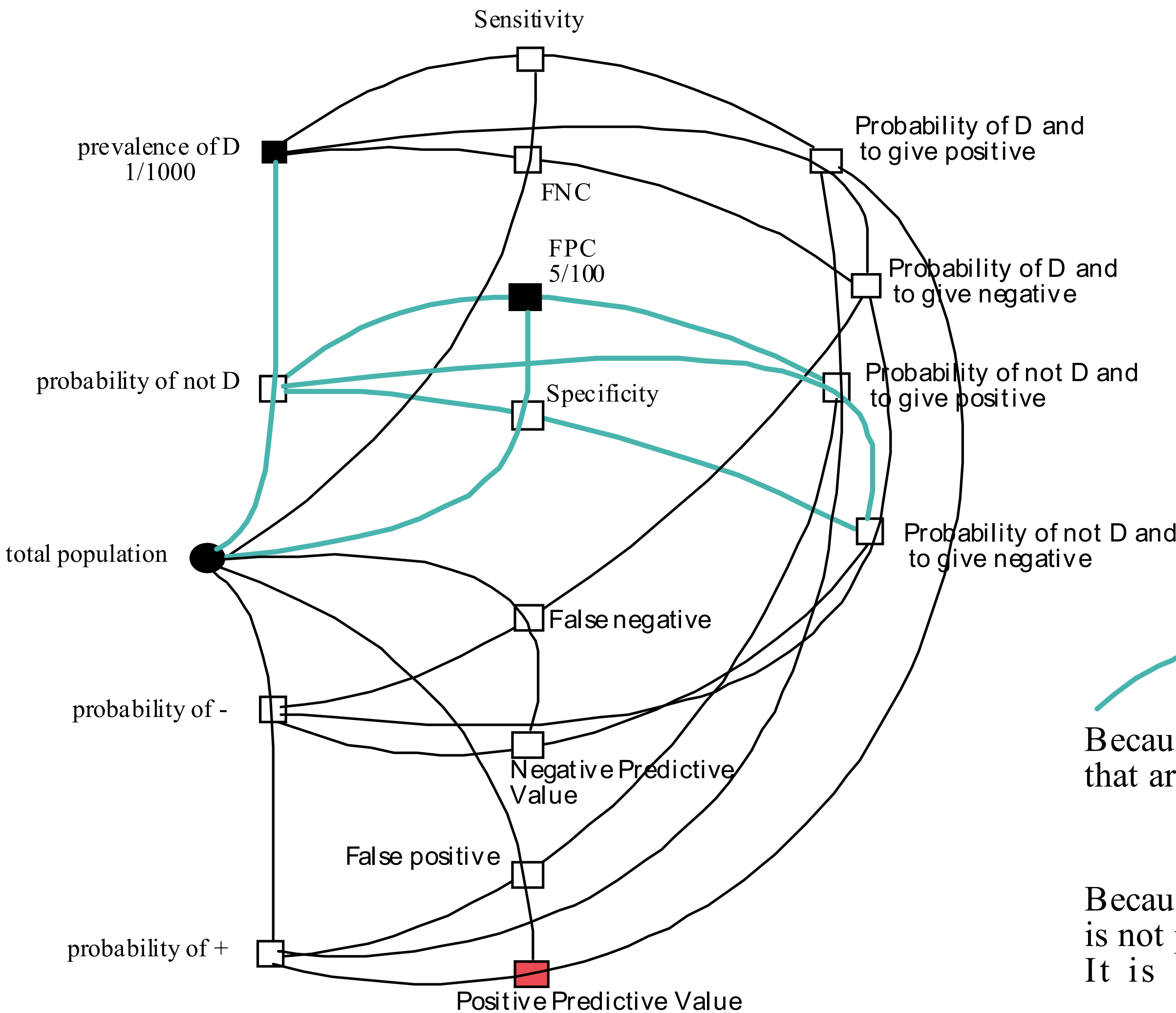
problems in context are analyzed following several aspects, like these:

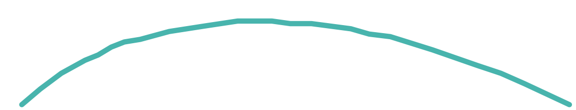
- context,
- phenomena (referring to events and probabilities),
- specific terminology (in this context),
- classification (according to their structure), and
- teaching setting or reference.

How does the trinomial graph work?

- If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have positive result actually has the disease, assuming you know nothing about the person’s symptoms or signs? (cited in Tversky & Kahneman, 1982, p. 154)
- 60 students and staff at Harvard Medical School
  - More than 30 participants gave 95% as the answer,
  - 11 participants answered 2% (but assuming the test correctly diagnoses every person who has the disease)

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about person's symptoms or signs.



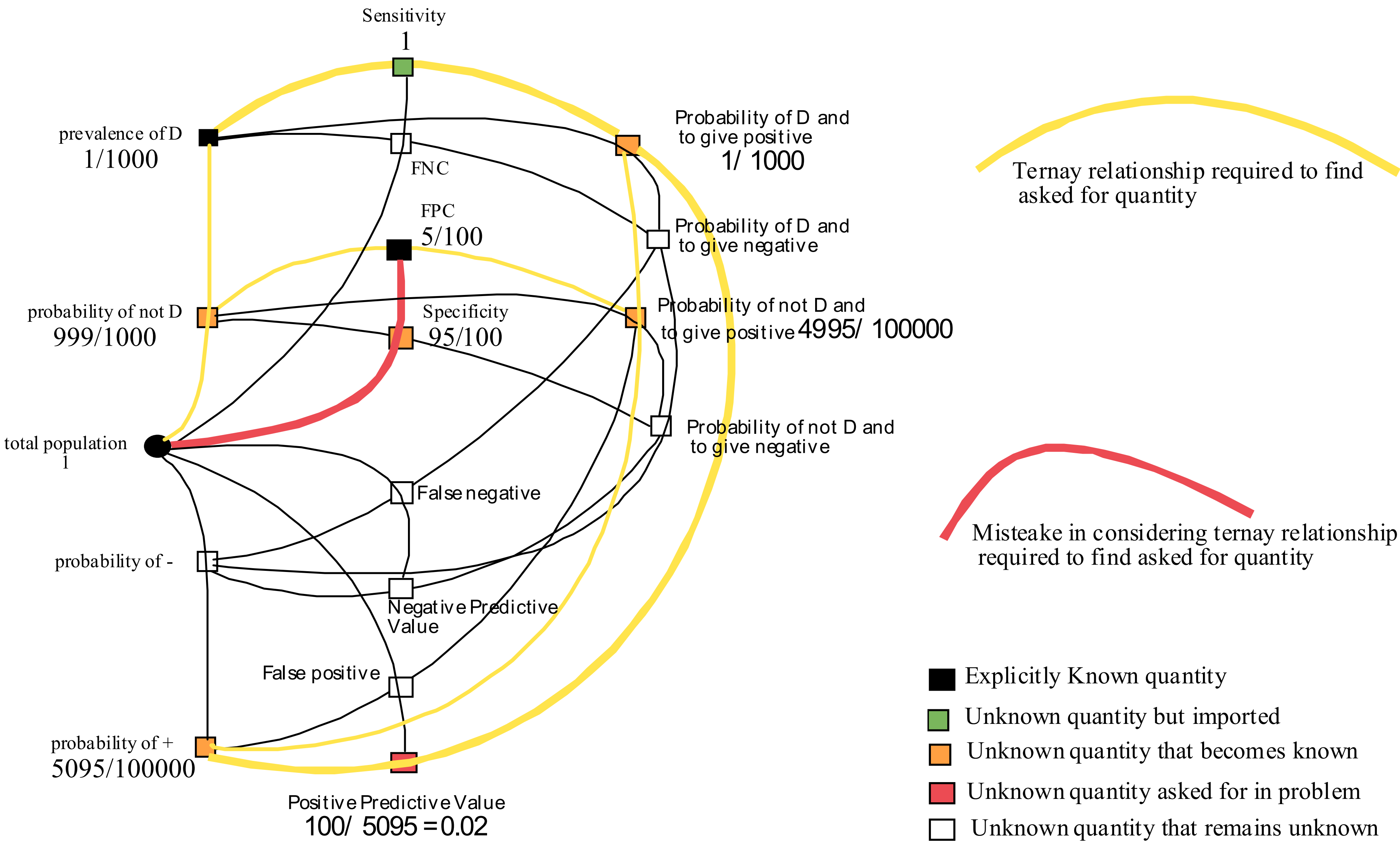


Because known data, ternary relashionships that are susceptible to be used

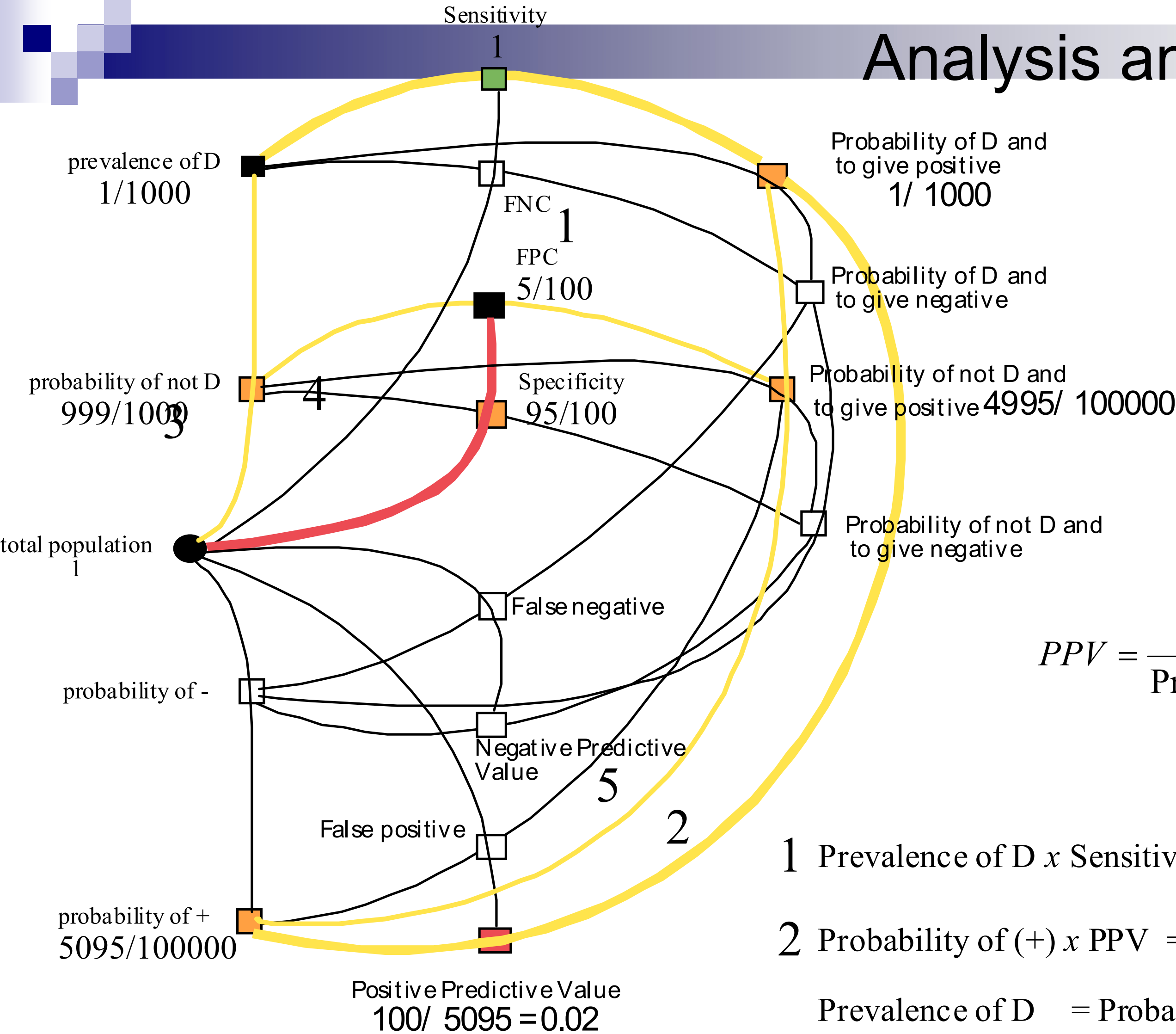
Because the trinomial graph is sub-dimensioned is not possible to find out the asked for quantity. It is still required one more known data.

11 participants answered 2% (assuming the test correctly diagnoses every person who has the dissease -sensitivity of the test is 1)

more than 30 participants gave 95% as the answer



Analysis and Synthesis method



Ternay relationship required to find asked for quantity

$$PPV = \frac{\Pr evalenceofD}{probabilityof(+)}$$

$$PPV = \frac{\Pr evalenceofD}{\Pr evalenceofD + \Pr obabilityofnotD \times FPC}$$

1 Prevalence of D *x* Sensitivity = Probability of (D *and* to give positive)

2 Probability of (+) *x* PPV = Probability of (D *and* to give positive)

Prevalence of D = Probability of (+) *x* PPV

3 4 Probability of not D *x* FPC = Probability of (not D *and* to give positive)

5 Probability of (D and to give positive) + Probability of (not D and to give positive) = Probability of (+)

Probability of (+) = Pevalence of D + Probability of not D *x* FPC

Synthesis

Of course, in order to answer the question only two known data are required !!!

Some results

- We have classified the world of ternary problems of conditional probability into four families and twenty subfamilies.
- Students' resolutions of problems of the N_2-family were investigated, analysing students' strategies of resolution, modes of functioning, mistakes and difficulties.
- Problems of the world of the ternary problems of conditional probability in Diagnostic test in Health has been investigated from a phenomenological approach.
- As a result, we have represented this world of problems in a trinomial graph.

Final comments

- We need to know more about problems in context of use. Probably, this knowledge let us to find answers to some questions such as: why?, for what?, and how? teaching conditional probability at schools.
- We also need to have a solid body of knowledge about how students solve these problems in context. What strategies they often use, what type of reasoning, what mistakes and difficulties they have …
- We already have some answers for problems of the N2-family, but they are not enough.

Thank you very much