# Combining Argumentation and Bayesian Nets for Breast Cancer Prognosis

Matt Williams* and Jon Williamson†

Progic, 7th July 2005

*Advanced Computation Laboratory, Cancer Research UK
†Department of Philosophy, Logic and Scientific Method, LSE

Two approaches to breast cancer prognosis:

- **Argumentation**: reasoning is clear but hard to weigh up
- **Statistical**: reliable conclusions, unclear chain of inference

▶ a combination might offer clarity and reliability

Two approaches to combining:

- **Integrated framework**: start from scratch
- **Modular framework**: off the shelf, simpler, customisable

Data + arguments $\longrightarrow$ statistical model $\longrightarrow$ prognosis + arguments.

**Plan**

- Breast cancer
- Logical formalism: argumentation
- Probabilistic formalism: causal nets
- Combination

## Breast Cancer

- **Very common**: commonest after skin cancer, 1 in 10, $\frac{1}{3}$ of all women's cancers.
- **Severe**: $\frac{1}{3}$ die.
- **Treatments**: surgery, radiotherapy, hormonal, chemotherapy.
- Reserve aggressive treatments for those at high risk of recurrence.

▶ Need to predict recurrence.

- This can be done on the basis of tumour characteristics, e.g. grade, size, number of involved lymph nodes.

## Logical Formalism: Argumentation

Early analysis and semantics by Dung (1995). Prakken and Sartor (1996) developed argument-based logic programming.

**Negation**:

- Classical negation $\neg L$,
- Negation-as-failure $\sim L$.

**Rules**:

- Strict (definitional) rules $HighMortality(x) \rightarrow \neg LowMortality(x)$,
- Defeasible rules $AggressiveTumour(x) \Rightarrow HighMortality(x)$.

**Arguments** are chains of instantiated rules, e.g.

- $AggressiveTumour(a)$
- $AggressiveTumour(a) \Rightarrow HighMortality(a)$
- $HighMortality(a) \rightarrow \neg LowMortality(a)$

**Attacking argument**:

- $Chemotherapy(a)$
- $Chemotherapy(a) \Rightarrow LowMortality(a)$

**Argumentation framework**: $< AR, attacks >$

- $AR$ is a set of arguments
- $attacks \subseteq AR \times AR$ is the attack relation.

**Plan**

- ✓ Breast cancer
- ✓ Logical formalism: argumentation
- ▶ Probabilistic formalism: causal nets
- • Combination

## Probabilistic Formalism: Causal Nets

**Epistemic causality**: what causal beliefs should an agent adopt, given her background knowledge?

- background knowledge: database, qualitative relationships
- database determines a probability distribution
- probabilistic dependencies are attributable to e.g. causal, logical, semantic connections
- causal beliefs should account for any unexplained strategic dependencies
- the agent should hold no unwarranted causal beliefs

Special case: no non-causal inducers of dependencies.

▶ causal graph is a minimal dag satisfying the Causal Markov Condition (each variable is probabilistically independent of its non-effects conditional on its direct causes)

▶ we can use standard algorithms to construct a causal belief graph.

**Causal net**:

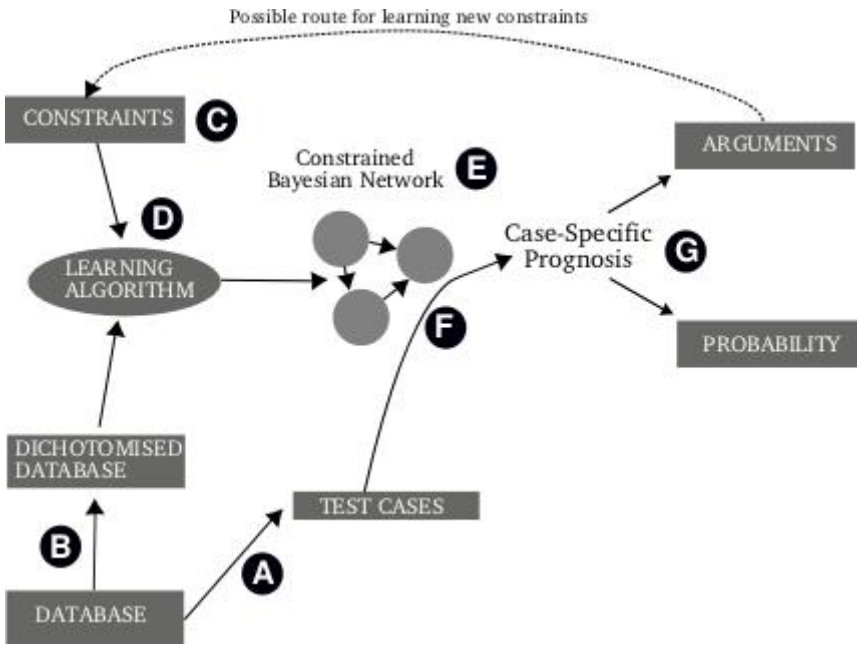
• By adding the probability distribution of each variable conditional on its direct causes we get a **causally-interpreted Bayesian net**

▶ we can calculate conditional probabilities.

A causal net can be used to:

- calculate prognostic probabilities
  e.g. $p(Recurrence \mid AggressiveTumour \wedge Chemotherapy) = 0.3$
- provide some justification for the prognosis
  e.g. $AggressiveTumour$ is a cause of $Recurrence$
  $Chemotherapy$ is a preventative of $Recurrence$

# Combining Argumentation and Causal Nets



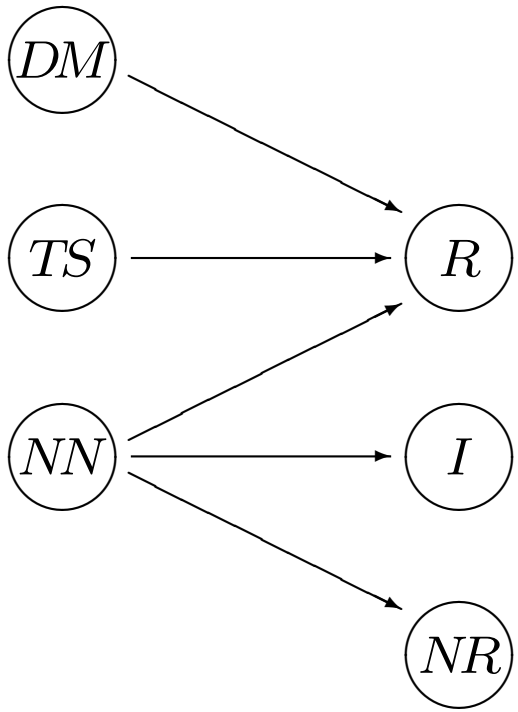
**Background knowledge**:

- Ljubljana Breast Cancer Dataset

Variables: age, menopause, tumour size, degree of malignancy, number of nodes, node capsule rupture, breast quadrant, breast, irradiation, recurrence.

- Qualitative constraints:
  - recurrence $\not\longrightarrow$ all,
  - irradiation $\not\longrightarrow$ all except recurrence.

No non-causal dependencies $\Rightarrow$ CMC $\Rightarrow$ standard algorithms:




CMC $\Rightarrow$ can use this graph as a basis for a Bayesian net,

- $p(R|A = 30-39, M = pre, TS = 35-59, DM = 3, NN = 0-2) = 0.2$.

Generate a set of arguments to support the prognosis:

- $TumourSize(a) > 29 \Rightarrow Recurrence(a)$
- $NumberOfNodes(a) \leq 2 \Rightarrow NonRecurrence(a)$
- $DegreeOfMalignancy(a) > 2 \Rightarrow Recurrence(a)$
- $NonRecurrence(a) \rightarrow \neg Recurrence(a)$

N.B.

- the net is used to generate the arguments.
- these arguments attack each other.
- other patient-specific arguments can be used to influence treatment decision.

**Plan**

✓ Breast cancer

✓ Logical formalism: argumentation

✓ Probabilistic formalism: causal nets

✓ Combination

**Further info**:

- written paper: progic web site
- Jon Williamson: **Bayesian nets and causality**, OUP 2005