# CAUSAL DETERMINATION IN ROAD ACCIDENTS: 

## An Application of Pearl's Probabilistic <br> Causal Models

Gary A. Davis<br>University of Minnesota

Tait Swenson<br>URS Corporation

Presented at:
Workshop on Investigating and Reporting Incidents and Accidents

> Williamsburg, VA
> September 17, 2003

Sponsored by:
Intelligent Transportation Systems Institute University of Minnesota

# WHAT DO WE MEAN BY A 'CAUSE' OF A ROAD ACCIDENT? 

Baker: "causal factor" is a circumstance "contributing to a result without which the result could not have occurred"

NTSB: "probable cause" is a "condition or event" such that "had the condition or event been prevented...the accident would not occur."

Tort Law: "cause in fact" "defendent's conduct is not a cause of the event, if the event would have occurred without it"

That is, we say 'A caused B' if
(1) A occurred,
(2) B occurred,
(3) if A hadn't occurred neither would B

> But how can we specify usable truth conditions for counterfactuals like (3)?

## RESEARCH ON SPEEDING AND ACCIDENT CAUSATION

To determine if an accident is caused by speeding, we need to answer three questions:
(A) What was the speed of the vehicle(s) involved?
(B) Was speeding a cause of this crash? That is, if the vehicle had been travelling at speed $v=v^{*}$, with $\mathrm{v}^{*}$ lower than the speed limit, would the crash have been avoided?
(C) And of course, what was the speed limit?

In principle, both (A) and (B) could be answered using deterministic accident reconstruction.

In practice, nontrivial uncertainties will be present, limiting us to uncertain answers to (A) and (B)

Our approach is to treat the accident reconstructionist as a Bayesian agent, so that the probability calculus can be used to reason about these uncertainties

## EXAMPLE VEHICLE/PEDESTRIAN ACCIDENT

## After-Accident Scene Map:



Synopsis:
7 year-old boy exited gate and attempted to cross road Driver braked to a stop leaving a $\mathbf{2 2} \mathbf{~ m}$. skidmark Speed limit= $30 \mathrm{mph}(48 \mathrm{~km} / \mathrm{h}, 13.4 \mathrm{~m} / \mathrm{s}$ )
Test skids performed after the accident yielded
$13.4 \mathrm{~m} / \mathrm{s}$
12.4 m

## PHILOSOPHICAL PRESUPPOSITIONS

If we are willing to accept:

## Local Laplacean Determinism:

In principle, it is possible to specify a set of structural equations, and a set of initial variable values, so that a given crash can be 'exactly' simulated (Baker, 1975)

## Counterfactual Treatment of Causation:

The causal effect of some variable is determined by comparing what happened to what would have happened, other things equal, had that variable been set at some different value.

## Probabilistic Treatment of Uncertainty:

"There is a forceful argument, that is being increasingly accepted, which concludes that the only sensible way to handle uncertainty is by means of probability." (D.V. Lindley, 1991)

## Then

Judea Pearl's theory of Probabilistic Causal Models can be used to pose and answer probabilistic versions of questions (A) and (B).

# GRAPHICAL MODEL OF A GENERIC ACCIDENT (Twin Network) 



Variables:
$\mathrm{V}=$ vehicle speed, $\mathrm{u}=$ other relevant factors
$\mathrm{z}=$ crash avoidance indicator,
$y=$ crash occurrence indictor
$\mathrm{e}=$ evidence
$\mathrm{v}^{*}=$ counterfactual speed, $\quad \mathrm{y}_{\mathrm{v}=\mathrm{v}^{*}}=$ counterfactual outcome

## PROBABILISTIC CAUSAL MODEL FOR A GENERIC CRASH

(1) Variables in model:

| $\mathrm{z}=$ | 1, crash avoidance situation arises <br> 0, no avoidance situation arises |
| :--- | :--- |
| $\mathrm{y}=$ | 1, crash occurs |
| 0, no crash occurs |  |
| $\mathrm{v}=$ | vehicle speed |
| $\mathrm{u}=$ | other relevant crash variables |
| $\mathrm{e}=$ | evidence concerning crash |

## (2) Deterministic core:

$$
\begin{array}{ll}
\mathrm{y}=\quad & 0, \text { if } \mathrm{z}=0 \text { or } \mathrm{f}(\mathrm{v}, \mathrm{u}) \geq 0 \\
& 1, \text { if } \mathrm{z}=1 \text { and } \mathrm{f}(\mathrm{v}, \mathrm{u})<0
\end{array}
$$

Example: Pedestrian steps in vehicle path ( $\mathrm{z}=1$ ) a distance x from an oncoming vehicle, and driver brakes to a stop:

$$
\mathrm{f}(\mathrm{v}, \mathrm{tp}, \mathrm{a}, \mathrm{x})=\mathrm{x}-\left(\mathrm{v}(\mathrm{tp})+\mathrm{v}^{2} / 2 \mathrm{a}\right)
$$

where $t p=$ reaction time, $\quad a=$ braking deceleration
(3) Evidence model: Evidence e is generated according to a distribution $P[e \mid y, v, u]$
(4) Prior Uncertainties: Specifying the reconstructionist's prior distributions on the exogenous variables $\mathrm{v}, \mathrm{z}, \mathrm{u}$ completes the probabilistic causal model.

## PROBABILISTIC CAUSAL CONCEPTS

Pearl defines a probabilistic version of "necessary cause," or "cause in fact," the "Probability of Necessity" (PN)
$P N=P\left[y_{v=v^{*}}=0 \mid y=1 \& v=v_{k}\right]$
The probability, other things equal, the crash would not have occurred had the vehicle's initial speed been $v^{*}$, given that the crash did occur and the initial speed was $\mathrm{v}_{\mathrm{k}}$

In practice, knowledge of vehicle's speed $\left(\mathrm{v}_{\mathrm{k}}\right)$ will be uncertain, so that probability of necessity cannot be applied as defined.

A more useful concept is what Pearl calls "Probability of Disablement," but which we will call the Probability of Avoidance (PA)
$\operatorname{PA}\left(v^{*}\right)=P\left[y_{v=v^{*}}=0 \mid y=1 \& e\right]$
$=\Sigma_{k} P\left[y_{v=v^{*}}=0 \mid y=1 \& v=v_{k}\right] P\left[v=v_{k} \mid y=1 \& e\right]$

This gives, the probability other things equal, that the crash would not have occurred had the vehicle been initially travelling at $\mathrm{v}^{*}$, given that the crash did occur, and the evidence at hand.

## METHODOLOGY

(1) $P\left[v=v_{i} \mid y=1 \& e\right]$ can be computed by an (in principle) straightforward application of Bayes theorem to the graphical model.
(2) Balke and Pearl's (1994) 3-step method can be used to compute the probabilities of counterfactual events:
(i) Abduction: compute $\mathrm{P}[\mathrm{u} \mid \mathrm{y}=1 \& \mathrm{e}]$
(ii) Action: set $v=v^{*}$
(iii) Prediction: compute

$$
\mathrm{PA}\left(\mathrm{v}^{*}\right)=\sum_{\left\{\mathrm{u}: \mathrm{y}\left(\mathrm{v}=\mathrm{v}^{*}\right)=0\right\}} \mathrm{P}[\mathrm{u} \mid \mathrm{y}=1 \& \mathrm{e}]
$$

(3) Balke and Pearl also show that the need to compute and store the posterior $\mathrm{P}[\mathrm{u} \mid \mathrm{y}=1 \& \mathrm{e}]$ can be avoided by applying Bayesian updating to a 'Twin Network' containing counterfactual variables.
(4) Computations can be carried out using Markov Chain Monte Carlo methods. In particular, the software BUGS and WINBUGS can be applied to this problem

## CASE 1: GREATRIX'S EXAMPLE ACCIDENT

## After-Accident Scene Map:



Synopsis:
7 year-old boy exited gate and attempted to cross road Driver braked to a stop leaving a $\mathbf{2 2} \mathbf{~ m}$. skidmark Speed limit= $30 \mathrm{mph}(48 \mathrm{~km} / \mathrm{h}, 13.4 \mathrm{~m} / \mathrm{s}$ )
Test skids performed after the accident yielded
$13.4 \mathrm{~m} / \mathrm{s}$
12.4 m

## COLLISION MODEL DETAILS



## Graphical Model:



Figure 1. Bayesian network.

## CASE 1 RESULTS

## Posterior Density for Speed, and Probability of Avoidance



## Conclusions:

(1) Speed limit was $48 \mathrm{~km} / \mathrm{h}$ ( 30 mph )
(2) Driver was quite probably speeding
(3) Speeding was quite probably a cause of this accident, in the sense that, other things equal, the accident would almost certainly have been avoided if the initial speed had been $48 \mathrm{~km} / \mathrm{h}$

# RESEARCH ON FREEWAY REARENDING ACCIDENTS 

Trajectories for 7 Vehicles, with Collision between 6 \& 7


## Brill's Rear-ending Collision Condition



## REAR-ENDING COLLISION MODEL

## Algebraic Collision Condition

$$
a_{k+1}<\left[v_{k+1}^{2}\right] /\left[\left(v_{k}^{2} / a_{k}\right)+2 v_{k+1}\left(h_{k+1}-r_{k+1}\right)\right]
$$

Hypothesis: Rear-end Collision Caused When Drivers' Reaction Times Exceed Their Following Headways.

Graphical Model for Three-Vehicle Platoon


## DETERMINING REACTION TIMES AND FOLLOWING HEADWAYS

(1) Speeds (v), braking decelerations (a) and braking initiation times ( t 0 ) estimated from observed trajectories
(2) Reaction times determined as difference between braking initiation times

(3) Following distance determined as location difference when lead vehicle initiated braking


## POSTERIOR MEANS AND STANDARD DEVIATIONS

|  | Vehicle |  |  |  |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $\mathrm{v}(\mathrm{fps})$ | 50.0 | 46.8 | 41.8 | 42.3 | 39.4 | 42.4 | 41.7 |
|  | $(0.76)$ | $(0.38)$ | $(0.39)$ | $(0.24)$ | $(0.24)$ | $(0.64)$ | $(0.39)$ |
| $\mathrm{a}\left(\mathrm{fps}^{2}\right)$ | 6.7 | 6.5 | 12.6 | 14.2 | 15.7 | 17.1 | 20.4 |
|  | $(0.11)$ | $(0.07)$ | $(0.89)$ | $(0.48)$ | $(0.87)$ | $(1.22)$ | $(1.09)$ |
| $\mathrm{amin}^{\left(\mathrm{fps}^{2}\right)}$ | -- | 6.2 | 11.4 | 12.7 | 14.1 | 16.9 | 24.5 |
| $\mathrm{~h}(\mathrm{sec})$ | -- | $1.05)$ | $(.63)$ | $(.38)$ | $(.56)$ | $(1.17)$ | $(1.64)$ |
| $\mathrm{r}(\mathrm{sec})$ | -- | 1.0 | 1.9 | 1.2 | 1.2 | 1.3 |  |
|  |  | $(.004)$ | $(.005)$ | $(.025)$ | $(.012)$ | $(.024)$ | $(.018)$ |
| $\mathrm{P}[\mathrm{r}>\mathrm{h}]$ | -- | .91 | 1.9 | 1.4 | 1.1 | 1.7 |  |
| $((.15)$ | $(.16)$ | $(.10)$ | $(.14)$ | $(.14)$ |  |  |  |

## Example Counterfactual Test:

If Driver 3's reaction time had been equal to his/her following headway would, other things equal, the crash have been prevented?

$$
\mathrm{PN}=\mathrm{P}\left[\operatorname{amin} 7_{\mathrm{r} 3<\mathrm{h} 3}<\mathrm{a}_{7}\right]=1.0
$$

## YES!

## CONCLUSIONS

If we are willing to accept:

## Local Laplacean Determinism:

In principle, it is possible to specify a set of structural equations, and a set of initial variable values, so that a given crash can be 'exactly' simulated (Baker, 1975)

## Counterfactual Treatment of Causation:

The causal effect of some variable is determined by comparing what happened to what would have happened, other things equal, had that variable been set at some different value.

## Probabilistic Treatment of Uncertainty:

"There is a forceful argument, that is being increasingly accepted, which concludes that the only sensible way to handle uncertainty is by means of probability." (D.V. Lindley, 1991)

## Then

Judea Pearl's theory of Probabilistic Causal Models can be used to pose and answer selected questions concerning causation in road accidents.

