Exploiting Innocuousness in Bayesian Networks

28th Australasian Joint Conference on Artificial Intelligence

Alexander Motzek* Ralf Möller*

* Universität zu Lübeck
Institute of Information Systems
Ratzeburger Allee 160, 23562 Lübeck, Germany
{motzek,moeller}@ifis.uni-luebeck.de

December, 4th 2015

Introduction

- (Dynamic) **Bayesian Networks**.
- New form of independence - **Innocuousness**.
- New form of DBNs - **Activator DBNs**.
- Formalize and Exploit.

Bayesian Networks

- **Syntactically** defined by a **graph** *B*.
- **Local semantics** as specifications of **local CPD**
- **Global semantic** as the **joint probability**

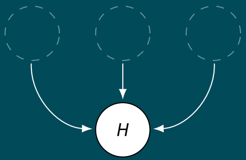
$$P(\vec{X}) = \prod_{X \in \vec{X}} P(X | parents(X))$$

Bayesian Networks

- **Graph** encodes guaranteed **independencies**.
- Not dependencies!
- Actual **dependencies** encoded/**specified in CPDs**.

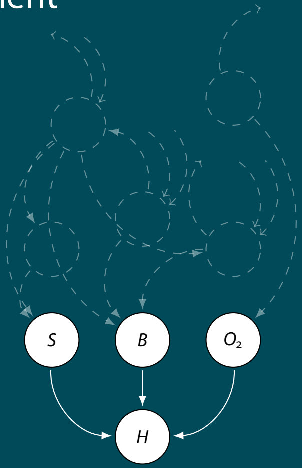
Independencies - A Gedankenexperiment

- Multiple **causes** can cause one **effect**.
- Our **hand is exposed** to various risks in a **blackbox**.
- **Exposures** can cause *H*arm.



Independencies - A Gedankenexperiment

- E.g., exposures to ***S*and**, ***B*unsen burner**, O_2
- $P(H|S, B, O_2)$
- $+s$and is present or not $\neg s$ present, a $+b$urner is turned on, $+o_2$ is present
- Exposures might be part of a much deeper probabilistic process.



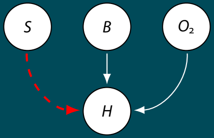
Causal Independence

- Classic. *S*and is **completely** irrelevant, i.e., **independent**.
- Investigated by Zhang and Poole, et al.

(i) change graph.

(ii) $P(H|S, B, O_2) = P(H|B, O_2)$ specifiable in a local CPD by

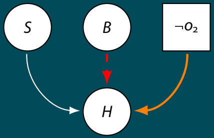
$$\forall h, b, o_2 : P(h| + s, b, o_2) = P(h|\neg s, b, o_2)$$



Context-Specific Independence

- A *B*unsen burner only works / can **only** causes harm **if** $+o_2$ **is present**.
- Investigated by Boutelier et al.
- $P(H|S, B, \neg o_2) = P(H|S, \neg o_2)$ specifiable in local CPD by

$$\forall s, h, b, o_2 : P(h|s, +b, \neg o_2) = P(h|s, \neg b, \neg o_2)$$
$$\exists s, h, b, o_2 : P(h|s, +b, +o_2) \neq P(h|s, \neg b, +o_2)$$

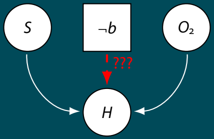


Innocuousness

- Allegedly, *Burner* is **only** relevant **if** it is **turned on** $+b$.
- A **turned off** $\neg b$urner is completely **irrelevant**, could have been left out.

- Very commonly found in Noisy-OR Assumptions.
  *"A **false** dependence **does not cause** any **harm**".*

- How to formalize?
  The relevant **context** $\neg b$ **is** the "ir**relevant**" random variable to be removed...
- $P(H|S, \neg b, O_2) = P(H|S, O_2)$ *not specifiable/expressible*?



Why formalize innocuousness?

Expressing Innocuousness is interesting:

- **More expressive** and causal specifications
- **Removing a link** is always good (computation time...)
- Can actually be **formalized** with and are **beneficial for ADBNs**. IJCAI'15

ADBNs - Running example

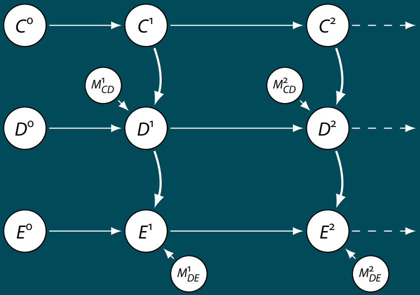
- **Regulatory compliance** of employees.
- A "creduluous" employee might **manipulate documents**.
- A credulous employee might (undeliberately) **influence other employees**.
- Might become credulous too, etc.

- Influences occur through **exchanged messages**.

- Track probabilistic credulousness-state over *time*.

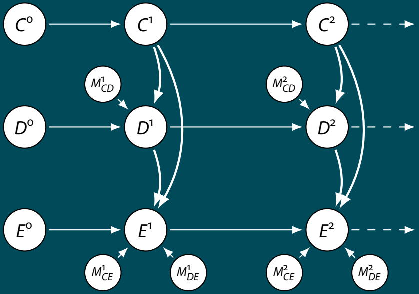
- Employees: ***C*laire, *D*on** and ***E*earl**.

Problem as a DBN



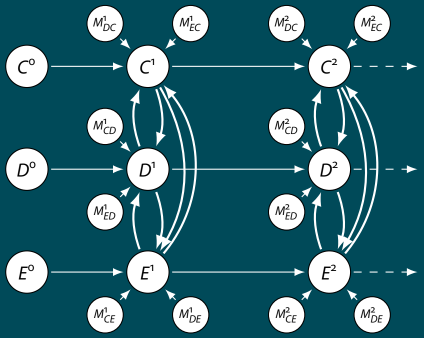
- Say, only *C*laire influences *D*on, influences *E*arl.
- i.e. *C* **influences** *E* **indirectly.**
- Typical DBN. ✓
- Problem correctly represented. ✓

Problem as a DBN



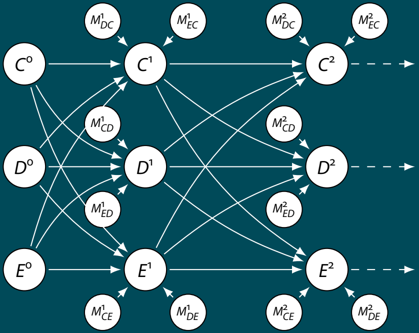
- Let's add some **more influences**.
- Claire *can* also influence Earl *directly*.
- Typical DBN. ✓
- Problem correctly represented. ✓

Problem as a DBN



- Say, **everybody** *can* influence **everybody**.
- "A BN is a **DAG**".
- Not a DBN. ✗
- Problem correctly represented. ✓?

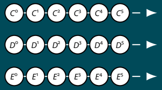
Problem as a DBN



- Resolve cycles over time.
- "Diagonal" inter-state dependencies.
- Common DBN . ✓
- Problem correctly represented. ✗ ?

DBN Restrictions

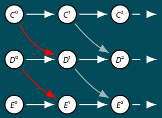
- "Diagonal" encodes "**incubation** time":
  t: Receive Message. t + 1*: Read and become influenced.*



a) Enforces **infinitesimal resolution of time** (e.g., seconds)

✗ High computation cost.

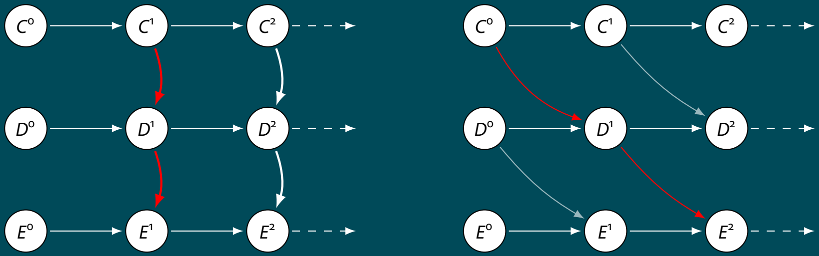
Observations not available this fine (e.g., only daily)?
Computation too costly? Transition only known hourly?



b) **Indirect influences not considerable**.

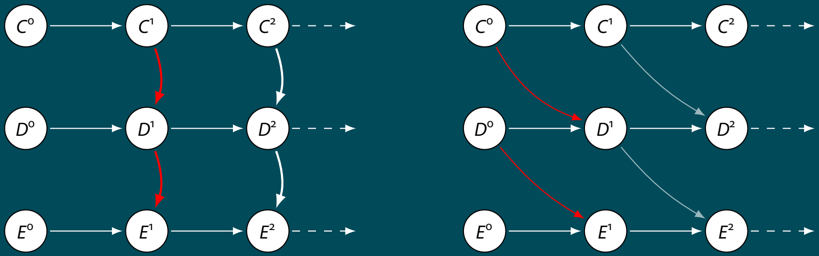
✗ Does not explain the world.

Classic DBNs spread indirect effects over time



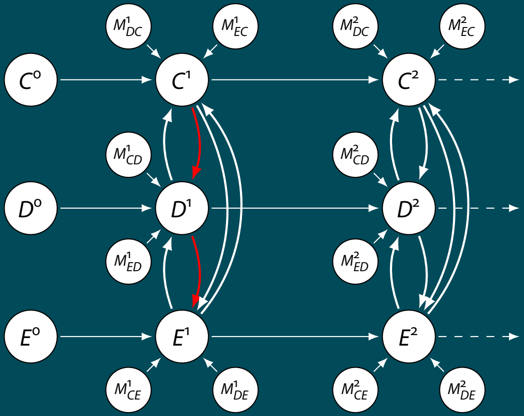
I.e., **observations** that require anticipations of **indirect effects** are **not supported**.

Classic DBNs spread indirect effects over time



I.e., **observations** that require anticipations of **indirect effects** are **not supported**.

Intuitive Design



Activator Random Variables

- Random variables M_{XY}^t representing **exchanged messages are special**
- M_{XY}^t have **activator nature**, i.e., are **Activator Random Variables**

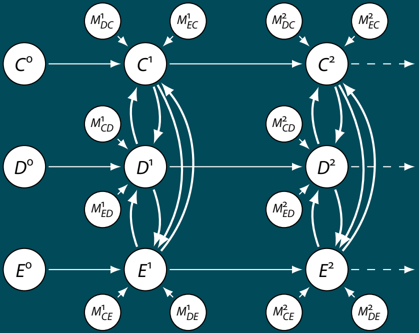
$$\forall x, x' \in \mathrm{dom}(X), \forall y \in \mathrm{dom}(Y), \forall \vec{z} \in \mathrm{dom}(\vec{Z}) :$$
$$P(y|x, \neg a_{XY}, \vec{z}) = P(y|x', \neg a_{XY}, \vec{z}) = P(y|*, \neg a_{XY}, \vec{z})$$

$*$ wildcard, $\vec{z}$ further dependencies

$$\exists x, x^* \in \mathrm{dom}(X), \exists y \in \mathrm{dom}(Y), \exists \vec{z} \in \mathrm{dom}(\vec{Z})$$
$$P(y|x, a_{XY}, \vec{z}) \neq P(y|x^*, a_{XY}, \vec{z})$$

- **Hint:** O_2 **is an activator for** B**unsen**

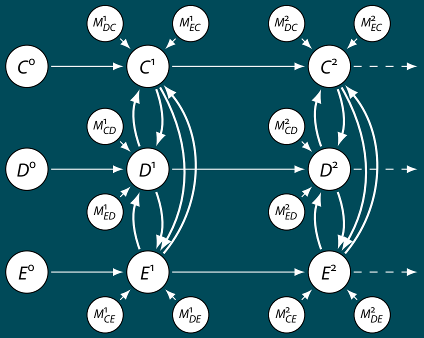
Activator Dynamic Bayesian Networks



- Is an *Activator Dynamic Bayesian Network*

- We show: **Semantically a (D)BN**, despite being based on a **cylic graph**!
- Straight forward semantic as joint probability as usual.

Activator Dynamic Bayesian Networks



- Is an *Activator Dynamic Bayesian Network*
- We show: **Semantically a (D)BN**, despite being based on a **cylic graph**!
- Straight forward semantic as joint probability as usual.
- Under some restrictions...

Restrictions Comparison

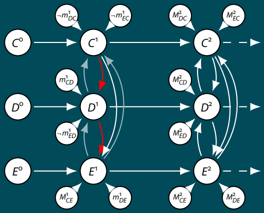
Cyclic ADBN

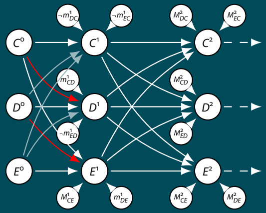
- No cyclic M^t_{XY} observations allowed.
- Activator set must form **DAG**.

"Diagonal" DBN

- No "interlocking" M^t_{XY} obs. allowed.
- must form **bipartite** graph.

#DAG >> #Bipartite



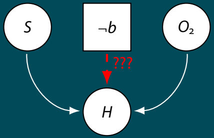


ADBN Contributions

- Bayesian networks can syntactically be based on **cyclic** graphs.
- Cyclic structures shall **not represent feedback-loops**.
- ADBNs are **well-defined**, if **activators are observed acyclic**.
- A required structure can not be known in advance and **is changing at every timestep**.

Formalizing Innocuousness in ADBNs - Recap

- Allegedly, *B***urner** is **only** relevant **if** it is turned **on** $+b$.
- Given $\neg b$, we could have removed the dependence.
- Problem: The **relevant context** $\neg b$ **is** the "ir**relevant**" random variable to be removed...



- in ADBNs $P(H|S, \neg b, O_2) = P(H|S, O_2)$ is now specifiable!

Formalizing Innocuousness in ADBNs

- O_2 is an **activator** for *B*unsen burner.
- Given $\neg o_2$, *B*urner becomes completely irrelevant. (Activator Nature)

- But also, given $\neg b$urner, O_2 becomes irrelevant.
- I.e., given $\neg b$urner, it is **like no oxygen** is present, i.e., *B*urner is irrelevant.

- This is **innocuousness**.
- We can actually **formalize** $P(H|S, \neg b, O_2) = P(H|S, O_2)$ by

$$\begin{aligned}\forall s, h, b, o_2 : P(h|s, \neg b, +o_2) &= P(h|s, \neg b, \neg o_2) \\ &= P(h|s, +b, \neg o_2\end{aligned}$$

Formalizing Innocuousness in ADBNs Formally

- $P(X|Y, A_{YX}, \vec{Z})$
- We say, A_{YX} can stand in multiple **innocuousness context** $\varphi_{A_{YX}} \in \vec{\varphi}_{A_{YX}}$
- This expresses, $P(X|\{\varphi_{A_{YX}} \backslash y\}, y, A_{YX}, \vec{Z}) = P(X|\{\varphi_{A_{YX}} \backslash y\}, A_{YX}, \vec{Z})$

Definition (Activator Innocuousness)

Let $\Phi_{A_{YX}}$ be the vector of random variables used in a context $\varphi_{A_{YX}}$ associated with A_{YX}. Every innocuousness context $\varphi_{A_{YX}} \in \vec{\varphi}_{A_{YX}}$ is then defined to hold

$$\forall x \in \mathrm{dom}(X), \forall \vec{z} \in \mathrm{dom}(\vec{Z}) : P(x|\varphi_{A_{YX}}, +a_{YX}, \vec{z}) = P(x|\varphi_{A_{YX}}, \neg a_{YX}, \vec{z}) \tag{1}$$

$$= P(x|\{\varphi_{A_{YX}} \backslash y \in \mathrm{dom}(Y)\}, y, \neg a_{YX}, \vec{z}) = P(x|\{\varphi_{A_{YX}} \backslash y\}, *, \neg a_{YX}, \vec{z}) , \tag{2}$$

with remaining arbitrary dependencies of X on other random variables $\vec{Z}$ and $\vec{z}$ as an arbitrary instantiation of those, excluding A_{YX} and $\Phi_{A_{YX}}$.

Exploiting Innocuousness in ADBNs

- In ADBNs, **Activators must "break** open **cycles"**.
- Only **DAG** observations were allowed.
- **Innocuousness** can formalize "a false dependent can be left out".
- Can **break open cycles**, too! (Proof see paper)
  Way more observations beyond DAGs allowed.

Exploiting Innocuousness in ADBNs - Example

- **Noisy OR** assumption for local CPDs in regulatory compliance domain.
- If employee C is compliant $\neg c$, he **does not influence** any other employees.
- Can "break open cycles".
- Refines **activator acyclicity constraint** $\mathfrak{A}$.
  Even cyclic message exchanges are well-defined.

ADBN and Innocuousness Contributions

- Bayesian networks can syntactically be based on **cyclic** graphs.
- **Well-defined beyond DAG** activator observations.
- **Innocuousness** properties **exist frequently**.
- Increase causal **expressiveness** of CPDs.

ADBNs provide

✓ **Free choice** of time granularity.
✓ More expressive CPD specifications.
✓ BN as world-representing first-class declaration.

Thank you

Fin and fin of conference.