

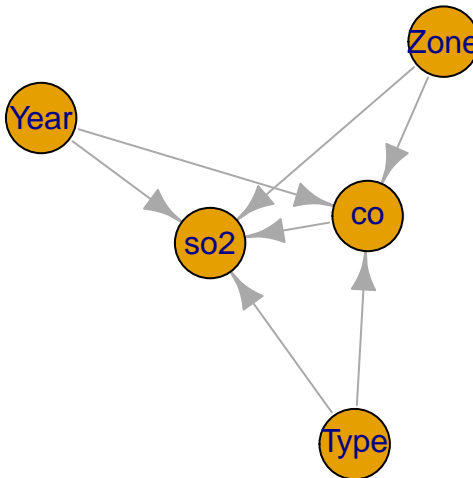
Causal-Inference_Do-Operator

Daniel Lindner, Simon Witzke

6/24/2020

Tests for do operator

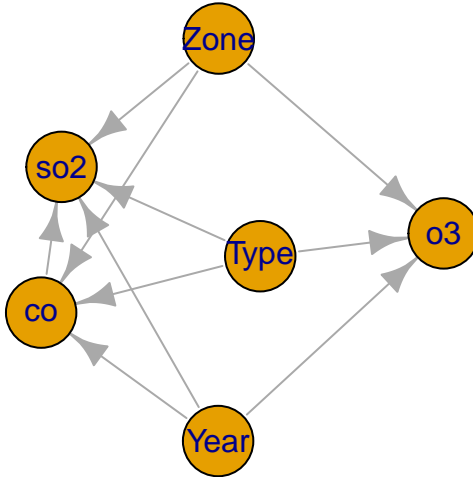
```
fig1 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -  
  +so2, Type - +so2, Year - +so2, co - +so2)  
plot.igraph(fig1, vertex.size = 35, vertex.label.family = "sans")
```



```
ce1 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig1, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(so2|Zone, Year, Type, co)P(Type)P(Year)P(Zone)$

```
fig2 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -  
  +so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -  
  +o3, Type - +o3)  
plot.igraph(fig2, vertex.size = 35, vertex.label.family = "sans")
```

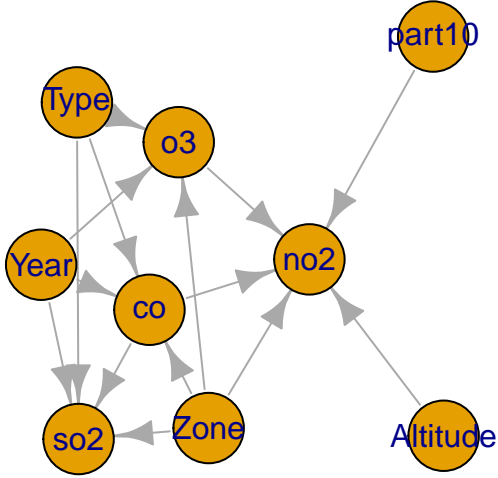


Example 2 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce2 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig2, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig3 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
  +so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
  +o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
  +no2, part10 - +no2)
plot.igraph(fig3, vertex.size = 35, vertex.label.family = "sans")
```

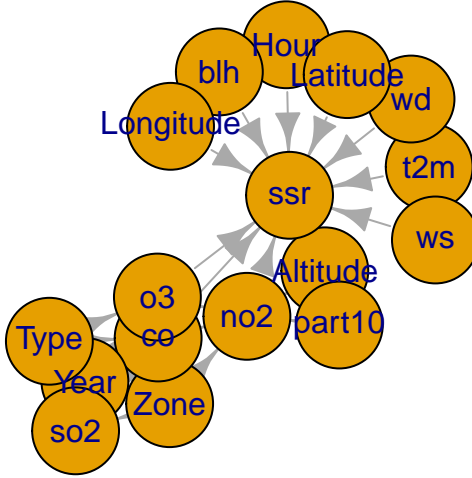


Example 3 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce3 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig3, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig4 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr)
plot.igraph(fig4, vertex.size = 45, vertex.label.family = "sans")
```

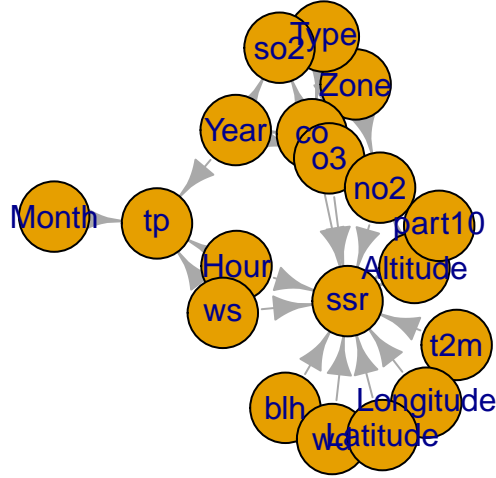


Example 4 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce4 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig4, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig5 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr, Month - +tp, ws - +tp, Year - +tp, Hour - +tp)
plot.igraph(fig5, vertex.size = 35, vertex.label.family = "sans")
```

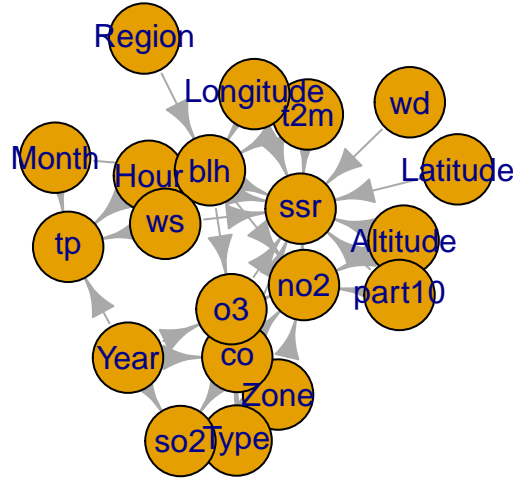


Example 5 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce5 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig5, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig6 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr, Month - +tp, ws - +tp, Year - +tp, Hour - +tp, blh -
+o3, blh - +no2, ws - +blh, Longitude - +blh, t2m - +blh,
Hour - +blh, Month - +blh, Region - +blh)
plot.igraph(fig6, vertex.size = 35, vertex.label.family = "sans")
```

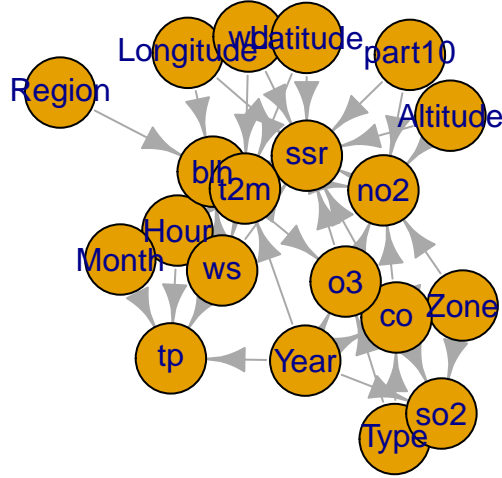


Example 6 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce6 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig6, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig7 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr, Month - +tp, ws - +tp, Year - +tp, Hour - +tp, blh -
+o3, blh - +no2, ws - +blh, Longitude - +blh, t2m - +blh,
Hour - +blh, Month - +blh, Region - +blh, t2m - +ws, Latitude -
+t2m, wd - +t2m, Month - +t2m, Year - +t2m, Hour - +t2m)
plot.igraph(fig7, vertex.size = 35, vertex.label.family = "sans")
```

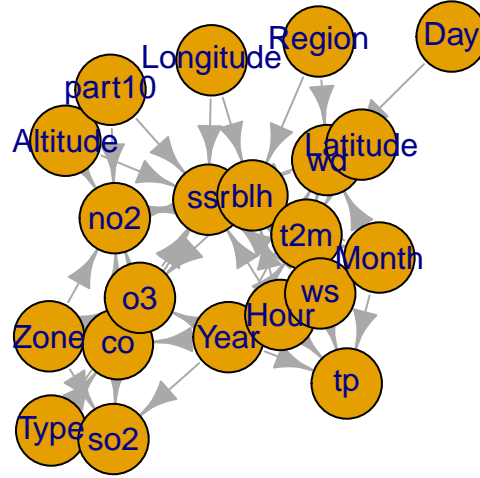


Example 7 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce7 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig7, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig8 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr, Month - +tp, ws - +tp, Year - +tp, Hour - +tp, blh -
+o3, blh - +no2, ws - +blh, Longitude - +blh, t2m - +blh,
Hour - +blh, Month - +blh, Region - +blh, t2m - +ws, Latitude -
+tp, wd - +t2m, Month - +t2m, Year - +t2m, Hour - +t2m,
Day - +wd, Month - +wd, Year - +wd, Region - +wd)
plot.igraph(fig8, vertex.size = 35, vertex.label.family = "sans")
```

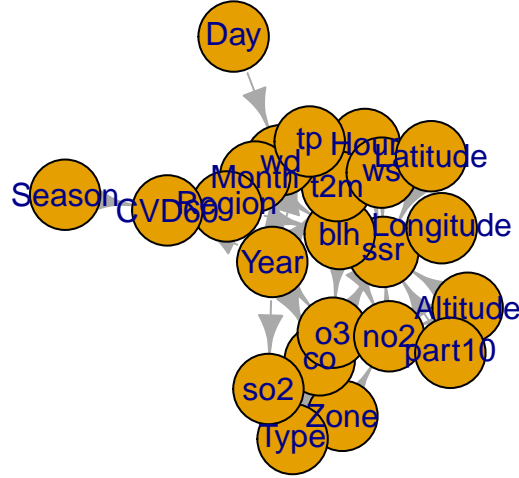


Example 8 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce8 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig8, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig9 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, part10 - +no2, co - +ssr, no2 - +ssr, part10 - +ssr,
o3 - +ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude -
+ssr, ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude -
+ssr, Month - +tp, ws - +tp, Year - +tp, Hour - +tp, blh -
+o3, blh - +no2, ws - +blh, Longitude - +blh, t2m - +blh,
Hour - +blh, Month - +blh, Region - +blh, t2m - +ws, Latitude -
+tp, wd - +t2m, Month - +t2m, Year - +t2m, Hour - +t2m,
Day - +wd, Month - +wd, Year - +wd, Region - +wd, Month - +CVD60,
Season - +CVD60, Year - +CVD60, Region - +CVD60)
plot.igraph(fig9, vertex.size = 35, vertex.label.family = "sans")
```

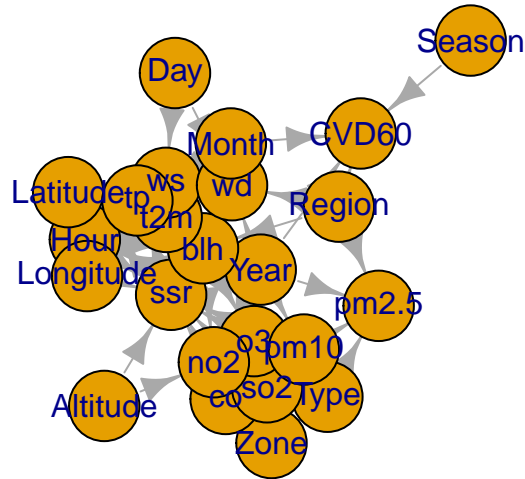



Example 9 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce9 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig9, expr = TRUE)
```

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

```
fig10 <- graph.formula(Zone - +co, Year - +co, Type - +co, Zone -
+so2, Type - +so2, Year - +so2, co - +so2, Zone - +o3, Year -
+o3, Type - +o3, Zone - +no2, o3 - +no2, co - +no2, Altitude -
+no2, pm10 - +no2, co - +ssr, no2 - +ssr, pm10 - +ssr, o3 -
+ssr, Hour - +ssr, blh - +ssr, t2m - +ssr, Longitude - +ssr,
ws - +ssr, wd - +ssr, Latitude - +ssr, Altitude - +ssr, Month -
+tp, ws - +tp, Year - +tp, Hour - +tp, blh - +o3, blh -
+no2, ws - +blh, Longitude - +blh, t2m - +blh, Hour - +blh,
Month - +blh, Region - +blh, t2m - +ws, Latitude - +t2m, wd -
+t2m, Month - +t2m, Year - +t2m, Hour - +t2m, Day - +wd,
Month - +wd, Year - +wd, Region - +wd, Month - +CVD60, Season -
+CVD60, Year - +CVD60, Region - +CVD60, pm10 - +pm2.5,
Type - +pm2.5, Year - +pm2.5, Region - +pm2.5, Year - +pm10,
Zone - +pm10, Type - +pm10, Day - +ws, Month - +ws, Latitude -
+ws, wd - +ws, Year - +ws, Longitude - +ws, Type - +no2,
Year - +no2, so2 - +ssr, ssr - +tp, wd - +blh)
plot.igraph(fig10, vertex.size = 35, vertex.label.family = "sans")
```



```
test <- igraph.from.graphNEL(bnlearn::as.graphNEL(bn))
```

```
difference(test, fig10)
```

```
## IGRAPH 9a5e434 DNW- 24 0 --
## + attr: name (v/c), weight (e/n)
## + edges from 9a5e434 (vertex names):
```

```
difference(fig10, test)
```

```
## IGRAPH 9a5edf1 DN-- 24 0 --
## + attr: name (v/c)
## + edges from 9a5edf1 (vertex names):
```

Example 10 do calculus: $P(\text{so2}|\text{do}(\text{co}))$

```
ce10 <- causal.effect(y = "so2", x = "co", z = NULL, G = fig10,
  expr = TRUE)
ceOriginal <- causal.effect(y = "so2", x = "co", z = NULL, G = test,
  expr = TRUE)
```

Result for manually created graph

$\sum_{Zone, Year, Type} P(\text{so2}|\text{Zone}, \text{Year}, \text{Type}, \text{co})P(\text{Type})P(\text{Year})P(\text{Zone})$

Result for “translated” graph

$P(so2|Zone, Type, Year, co)$

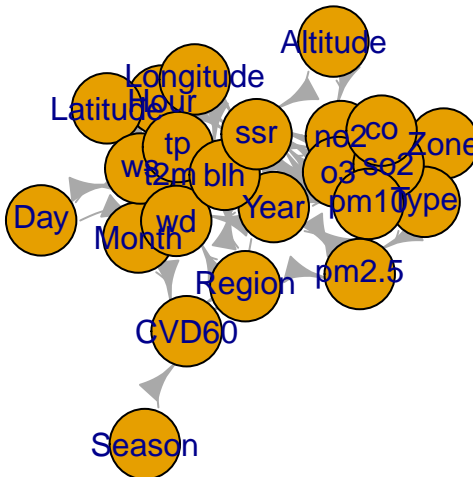
```
test <- igraph.from.graphNEL(bnlearn::as.graphNEL(bn))

tkplot(test, vertex.size = 35, vertex.label.family = "sans", vertex.color = "white")

eff <- causal.effect(y = "so2", x = "co", z = c("o3"), test, expr = T)

cat(eff)

graph <- igraph.from.graphNEL(bnlearn::as.graphNEL(bn))
plot.igraph(graph, vertex.size = 35, vertex.label.family = "sans",
  edge.curved = T)
```



Example do calculus: Complex

```
eff <- causal.effect(y = "so2", x = "co", z = c("Year", "Type",
  "Zone"), graph, expr = T)
```

$$\left(\frac{P(so2|Zone, Type, Year, co)}{\sum_{so2} P(so2|Zone, Type, Year, co)} \right),$$

```
eff2 <- causal.effect(y = "so2", x = "co", z = names(bn)[!names(bn) %in%
  c("so2", "co")], graph, expr = T)
```

$$\left(\frac{P(ssr|Region, Zone, Type, Year, Month, Day, Hour, Latitude, Longitude, Altitude, co, pm10, wd, so2, t2m, ws, blh, o3, no2)P(so2|Zone, Type, Year, co)}{\sum_{so2} P(ssr|Region, Zone, Type, Year, Month, Day, Hour, Latitude, Longitude, Altitude, co, pm10, wd, so2, t2m, ws, blh, o3, no2)P(so2|Zone, Type, Year, co)} \right),$$