Logique & Analyse **185–188** (2004), 33–51

APPLICATIONS OF THE ADAPTIVE LOGIC FOR CAUSAL DISCOVERY*

LEEN DE VREESE AND ERIK WEBER

1. *Introduction*

In his article *Causal Discovery Using Adaptive Logics*, Maarten Van Dyck proposed a logic that allows one to derive causal statements from probabilistic information. This causal logic is adaptive. One of his purposes in using this sort of logic is trying to get a grip on how *humans* reason from association to causation. All adaptive logics indeed serve the common goal of formalizing real life reasoning processes. Our purpose in this article is trying to demonstrate that Van Dyck's adaptive logic for causal discovery (henceforth ALcd) does reach this goal quite well, by providing two illustrations from the field of social sciences. A first and short example will be given in section 2. A more elaborate illustration can be found in section 3. Those applications will give us better insight in possible extensions of the logic that will strengthen its applicability. This will be discussed in the concluding section 4.

For the readers' convenience, we put a list with all the rules of ALcd here. We refer to [4] for an extensive description of their function and meaning.

A1 $\neg\mathcal{P}(\alpha,\alpha)$

A2 $\neg(\alpha \amalg \beta) \equiv (\alpha \rightarrow \beta \vee \beta \rightarrow \alpha \vee DSO_{\alpha\beta} \vee \alpha = \beta)$

A2' $\neg(\alpha \amalg \beta) \equiv (\mathcal{P}(\alpha,\beta) \vee \mathcal{P}(\beta,\alpha) \vee (\exists\gamma)(\mathcal{P}(\gamma,\alpha) \wedge \mathcal{P}(\gamma,\beta)) \vee \alpha = \beta)$

*We like to thank some people for their help: Farah Focquaert for the search for the examples; Maarten van Dyck and referees Libor Behounek and Adam Grobler for their useful comments. The research for this paper benefited from a bilateral scientific exchange project funded by the Ministry of the Flemish Community (project BIL/01/80) and the Polish State Committee for Scientific Research (KBN).

R1 $\dfrac{DSO_{\alpha\beta}, \quad [\mathcal{P}(\alpha,\gamma) \wedge \mathcal{P}(\gamma,\beta)] \vee [\mathcal{P}(\beta,\gamma) \wedge \mathcal{P}(\gamma,\alpha)] \vee [\mathcal{P}(\gamma,\alpha) \wedge \mathcal{P}(\gamma,\beta)] \vee (\exists\delta)[\mathcal{P}(\delta,\gamma) \wedge \mathcal{P}(\gamma,\alpha) \wedge \mathcal{P}(\delta,\beta)] \vee (\exists\delta)[\mathcal{P}(\delta,\alpha) \wedge \mathcal{P}(\delta,\gamma) \wedge \mathcal{P}(\gamma,\beta)]}{\gamma \in \Psi(\mathcal{SO}_{\alpha\beta})}$

R2 $\dfrac{\gamma \in \Psi(\mathcal{SO}_{\alpha\beta})}{\neg(\alpha \rightarrow \gamma \wedge \beta \rightarrow \gamma)}$

R3 $\dfrac{\neg(\alpha \amalg \beta), \quad \gamma \in \Psi(\mathcal{SO}_{\alpha\beta})}{[\mathcal{P}(\alpha,\gamma) \wedge \mathcal{P}(\gamma,\beta)] \vee [\mathcal{P}(\beta,\gamma) \wedge \mathcal{P}(\gamma,\alpha)] \vee [\mathcal{P}(\gamma,\alpha) \wedge \mathcal{P}(\gamma,\beta)] \vee (\exists\delta)[\mathcal{P}(\delta,\gamma) \wedge \mathcal{P}(\gamma,\alpha) \wedge \mathcal{P}(\delta,\beta)] \vee (\exists\delta)[\mathcal{P}(\delta,\alpha) \wedge \mathcal{P}(\delta,\gamma) \wedge \mathcal{P}(\gamma,\beta)]}$

DR1 $\dfrac{\alpha \neq \beta, \quad (\alpha \amalg \beta), \quad \alpha \rightarrow \gamma \vee \gamma \rightarrow \alpha \vee DSO_{\alpha\gamma}, \quad \beta \rightarrow \gamma \vee \gamma \rightarrow \beta \vee DSO_{\beta\gamma}}{\alpha \rightarrow \gamma \vee DSO_{\alpha\gamma}, \quad \beta \rightarrow \gamma \vee DSO_{\beta\gamma}}$

DR2 $\dfrac{\alpha \neq \beta, \quad DSO_{\alpha\beta}, \quad \gamma \notin \Psi(\mathcal{SO}_{\alpha\beta}), \quad \alpha \rightarrow \gamma \vee \gamma \rightarrow \alpha \vee DSO_{\alpha\gamma}, \quad \beta \rightarrow \gamma \vee \gamma \rightarrow \beta \vee DSO_{\beta\gamma}}{\alpha \rightarrow \gamma \vee DSO_{\alpha\gamma}, \quad \beta \rightarrow \gamma \vee DSO_{\beta\alpha}}$

Rc $\dfrac{\neg(\alpha \amalg \beta)}{\alpha \rightarrow \beta \vee \beta \rightarrow \alpha \vee \alpha = \beta \vee DSO_{\alpha\beta}}$

As a rule for the marking of lines we have:
Where Θ is the fifth element of line i, line i is marked iff a formula δ is unconditionally derived for some $\delta \in \Theta$.

2. *Boot Camps and Recidivism*

As a first and simple application of ALcd we use the formalization of a hypothetical study presented by Schutt [7]. The hypothetical study investigates

the relationship between attending a boot camp (B) during imprisonment and the likelihood of recidivism (R) after release. From the fact that there is an association between boot camps and recidivism, we might infer that there is a causal relation. This can be formalized as follows:

1	$\neg(B = R)$	–	PREM	$\emptyset$
2	$\neg(B \amalg R)$	–	PREM	$\emptyset$
3	$(B \rightarrow R) \vee (R \rightarrow B) \vee (B = R)$	2	Rc	$\mathcal{SO}_{B\,R}\sqrt{}_{11}$
4	$(B \rightarrow R) \vee (R \rightarrow B)$	1,3	CL	$\mathcal{SO}_{B\,R}\sqrt{}_{11}$
5	$\neg(R \rightarrow B)$	–	BK	$\emptyset$
6	$B \rightarrow R$	4,5	CL	$\mathcal{SO}_{B\,R}\sqrt{}_{11}$

The BK-line introduces some background-knowledge: time order excludes the causal relation in one of the directions. In this particular case line 5 explicates that we reason from the background knowledge that the boot camp comes first, and is possibly followed by recidivism after release.

Nonetheless, when statistically controlling on gender (G), it might turn out that the positive correlation between boot campers and recidivism disappears. The apparent evidence might then be explained by the fact that women are more likely to attend boot camp and are less likely to commit crimes after prison, no matter if they attended boot camp or not. The continuation of the reasoning process with those characteristics can be formalized as follows:

7	$\neg(B = G)$	–	PREM	$\emptyset$
8	$\neg(G = R)$	–	PREM	$\emptyset$
9	$\neg(G \amalg B)$	–	PREM	$\emptyset$
10	$\neg(G \amalg R)$	–	PREM	$\emptyset$
11	$(B \amalg R \mid G)$	–	PREM	$\emptyset$

The introduction of a new premise on line 11 obligates us to mark line 3, 4 and 6. On the other hand, the new premises provide opportunities to continue our reasoning process:

12	$(G \rightarrow B) \vee (B \rightarrow G) \vee (B = G)$	9	Rc	$\mathcal{SO}_{B\,G}$
13	$(G \rightarrow B) \vee (B \rightarrow G)$	7,12	CL	$\mathcal{SO}_{B\,G}$
14	$(G \rightarrow R) \vee (R \rightarrow G) \vee (R = G)$	10	Rc	$\mathcal{SO}_{R\,G}$
15	$(G \rightarrow R) \vee (R \rightarrow G)$	8,14	CL	$\mathcal{SO}_{R\,G}$
16	$\neg(B \rightarrow G)$	–	BK	$\emptyset$
17	$\neg(R \rightarrow G)$	–	BK	$\emptyset$
18	$G \rightarrow B$	13,16	CL	$\mathcal{SO}_{B\,G}$
19	$G \rightarrow R$	15,17	CL	$\mathcal{SO}_{R\,G}$

It is evident we reason from the fact that neither boot camps, nor recidivism will have any effect on gender, in accordance with the time order underlying these variables. This justifies BK-lines number 16 and 17.

3. *Parental Separation and Smoking Initiation*

3.1. *Introduction*

Let's turn to another example of a real investigation. In the previous section, we have been using ALcd to formalize the reasoning process that leads from the premises to the results. The following example will not only be more elaborate, but one that illustrates meanwhile that the same logic can also clarify the heuristic process of researchers. I.e.: What leads the scientist to the specific questions essential for the good proceeding and appropriate results of the study?

James Kirby (2002) made a study in which he tried to find out whether there is an influence of parental separation on smoking initiation in adolescents. While a lot of research finds a positive association between family disruption and smoking, Kirby wanted to go one step further. He also was searching for the causal pathway of this connection. For the precise outline and results of his study we refer to [5]. What is important here, is that we can formalize the reasoning process and findings of Kirby very well, using ALcd. Formalizing this study will give us an elaborate example of the application of ALcd within the social sciences. In a first phase of the survey, the central question was if there is a relationship between parental separation and the likelihood that an adolescent will start smoking which was not due to (a) common cause(s) of both. In this part of Kirby's reasoning process we will use ALcd as clarifying for the heuristic process followed by Kirby. For this to be possible, we will have to introduce one more conditional rule and one more marking rule.

First, at any stage in the proof, we permit the introduction of a question $?[\alpha]$ on the condition $(\alpha, \neg\alpha)$. We will call this rule QRc. This rule will be combined with a marking rule for questions, which states that a line with a question $?[\alpha]$ can be marked, whenever α or $\neg\alpha$ has been derived:

QRc At any stage in the proof, one may add to the proof a line consisting of:
(i) the appropriate line number
(ii) $?[\alpha]$
(iii) -

(iv) QRc
(v) $(\alpha, \neg\alpha)$

Definition 1: (Marked Line for Questions) *A line i that has* $(\alpha, \neg\alpha)$ *as its fifth element is marked at stage s of a proof iff at that stage either* α *or* $\neg\alpha$ *is introduced.*

A second phase of the survey concentrates on the possible causal pathways through which parental separation has an influence on smoking initiation. Here, ALcd will not only be used for the characterization of the heuristic process, but also as in the previous example for the characterization of the reasoning from data to results. The two research questions form two parts in the reasoning process which will be analysed separately, respectively in section 3.2 and 3.3.

3.2. *Controlling for Common Causes*

The first question in the research of Kirby, namely whether there is a relationship between parental separation and the likelihood that an adolescent will start smoking, was quite easy to answer. The results of the statistical analysis clearly showed an association between parental separation and smoking initiation. This can be formalized as in line 2.

1	$\neg(PS = SI)$	–	PREM	$\emptyset$
2	$\neg(PS \amalg SI)$	–	PREM	$\emptyset$
3	$(PS \rightarrow SI) \vee (SI \rightarrow PS) \vee (PS = SI)$	2	Rc	$\mathcal{SO}_{PSSI}$
4	$(PS \rightarrow SI) \vee (SI \rightarrow PS)$	1,3	CL	$\mathcal{SO}_{PSSI}$
5	$\neg(SI \rightarrow PS)$	–	BK	$\emptyset$
6	$PS \rightarrow SI$	4,5	CL	$\mathcal{SO}_{PSSI}$

Some clarifications about the continuation of the proof in line 3 till 6. Line 3 and 4 result from the application of the logical rules. In line 5 some background knowledge has been introduced. We infer this knowledge from the time order implemented in Kirby's study. In the concrete Kirby examines if the experience of the divorce of the parents between time wave 1 and time wave 2 resulted in smoking initiation at time wave 2. This set-up of the study has been taken as a justification for the introduction of the background knowledge as implemented on line 5. The application of classical logic leads then to line 6, which states that there is indeed a causal path from parental separation to smoking initiation.

The next step was to examine if this association persisted after controlling for various confounding variables. In other words: was the association between parental separation and smoking initiation not due to other factors

associated with both? Kirby was looking here for possible common causes of both parental separation and smoking initiation, which could make havoc of the results if they were not identified as such beforehand. Or translated into the logical proof: Kirby was searching for reasons to mark line 3, 4 and 6. Previous research offered him some background knowledge about possible confounding variables. We limit ourselves here to those variables which are typical for this study. In other words we skip the control variables that recur in each statistical study, namely age, gender and race.[1] We will use the abbreviation HI for household income, PE for parental education, and SP for smoking status of parents:

> *Household income has been found to be negatively associated with the likelihood of divorce, and it is also likely to be associated with smoking initiation in children. Therefore, household income at wave 1 is included as a control variable.* [5, p. 61]

Or in formalized form:

7	$PS \neq HI$	–	PREM	$\emptyset$
8	$HI \neq SI$	–	PREM	$\emptyset$
9	$\neg(PS \amalg HI)$	–	PREM	$\emptyset$
10	$\neg(HI \amalg SI)$	–	PREM	$\emptyset$

Given the consequences of line 2, 9 and 10, as written respectively on line 4, 12 and 14, the possibility arises that household income is a common cause of both parental separation and smoking initiation which possibly screens them off (cf line 19). This would implicate that the possibility of a causal path between parental separation and smoking initiation (cf line 6) disappears when controlling on household income. By consequence, this reasoning leads to the research question on line 20.

11	$PS \rightarrow HI \vee HI \rightarrow PS \vee PS = HI$	9	Rc	$\mathcal{SO}_{PS\,HI}$
12	$PS \rightarrow HI \vee HI \rightarrow PS$	7,11	CL	$\mathcal{SO}_{PS\,HI}$
13	$HI \rightarrow SI \vee SI \rightarrow HI \vee SI = HI$	10	Rc	$\mathcal{SO}_{HI\,SI}$
14	$SI \rightarrow HI \vee HI \rightarrow SI$	8, 13	CL	$\mathcal{SO}_{HI\,SI}$
15	$\neg(PS \rightarrow HI)$	–	BK	$\emptyset$
16	$\neg(SI \rightarrow HI)$	–	BK	$\emptyset$
17	$HI \rightarrow PS$	12, 15	CL	$\mathcal{SO}_{PS\,HI}$
18	$HI \rightarrow SI$	14, 16	CL	$\mathcal{SO}_{HI\,SI}$
19	$HI \rightarrow SI \wedge HI \rightarrow PS$	17,18	CL	$\mathcal{SO}_{PS\,HI} \cup \mathcal{SO}_{HI\,SI}$
20	$?[(PS \amalg SI \mid HI)]$	-	QRc	

[1] For the bibliographical notes concerning these background knowledge, we refer to the original article of Kirby himself. [5]

		$(PS \amalg SI\|HI), \neg(PS \amalg SI\|HI)$	$\surd$	
21	$\neg(PS \amalg SI\|HI)$	-	PREM	$\emptyset$

On which grounds BK-lines 15 and 16 are added? It is clear from the upset of Kirby's study that he included the household income at time wave 1, cases of divorce between time wave 1 and 2 and smoking initiation at time wave 2. This implies that smoking initiation can only be an effect-variable of household income and parental separation, and no cause-variable.

When statistically controlling on household income, Kirby nonetheless didn't find reasons to drop the causal path between parental separation and smoking initiation (cf line 21). An association between parental separation and smoking initiation stays present, even when controlling on household income. Therefore, line 20 can be marked and the marking of line 3, 4 and 6 should not be executed. This implies further the possibility of an indirect effect of household income on smoking initiation (when $HI \rightarrow PS$ and $PS \rightarrow SI$). By consequence this variable is introduced as a control variable.

A second possible confounding variable is the following one:

> *Parental education is another variable that could be a source of bias if omitted from my analysis. People with less education are more likely to get divorced than people with more education. Furthermore, children of parents with less education are more likely to smoke than children with more educated parents. If parental education is not included in my analysis, some of the influence that education has on smoking initiation may be attributed to parental separation.* [5, p. 61]

Or formalized:

22	$PS \neq PE$	–	PREM	$\emptyset$
23	$PE \neq SI$	–	PREM	$\emptyset$
24	$\neg(PS \amalg PE)$	–	PREM	$\emptyset$
25	$\neg(PE \amalg SI)$	–	PREM	$\emptyset$

An analogous line of reasoning as with regard to household income has been made with regard to parental education (PE). The consequences of the premises as written on line 4, 27 and 29 result in the possibility that parental education is a common cause of parental separation and smoking initiation. The causal path between those latter two variables (cf line 3, 4 and 6) would then disappear when controlling on parental education. All this leads us to the research question on line 35.

26	$PS \rightarrow PE \vee PE \rightarrow PS \vee PS = PE$	24	Rc	$\mathcal{SO}_{PS\,PE}$
27	$PS \rightarrow PE \vee PE \rightarrow PS$	22, 26	CL	$\mathcal{SO}_{PS\,PE}$
28	$PE \rightarrow SI \vee SI \rightarrow PE \vee PE = SI$	25	Rc	$\mathcal{SO}_{PE\,SI}$

29	$PE \rightarrow SI \vee SI \rightarrow PE$	23, 28	CL	$\mathcal{SO}_{PE\,SI}$
30	$\neg(PS \rightarrow PE)$	–	BK	$\emptyset$
31	$\neg(SI \rightarrow PE)$	–	BK	$\emptyset$
32	$PE \rightarrow PS$	27, 30	CL	$\mathcal{SO}_{PS\,PE}$
33	$PE \rightarrow SI$	29, 31	CL	$\mathcal{SO}_{PE\,SI}$
34	$PE \rightarrow SI \wedge PE \rightarrow PS$	32,33	CL	$\mathcal{SO}_{PS\,PE} \cup \mathcal{SO}_{PE\,SI}$
35	$?[(PS \amalg SI \mid PE)]$	-	QRc	$(PS \amalg SI \mid PE), \neg(PS \amalg SI \mid PE)$ $\surd$
36	$\neg(PS \amalg SI \mid PE)$	-	PREM	$\emptyset$

Normally, parental education comes first in order of time, before parental separation and smoking initiation of children. This justifies BK-lines 30 and 31.

Again, no reason to accept the proposal of line 35 has been found. By consequence, there's again no motive to cast doubt on the association between parental separation and smoking initiation (cf line 36). Line 35 can be marked, and the marking of line 3, 4 and 6 should not be carried out.

A last possible confounding variable is the smoking status of parents (SP):

> *Another potential confounding variable is the smoking status of parents. ..., parents provide models of behavior for their children. If a parent smokes, it is likely that a child will see smoking as more acceptable than if a parent does not smoke. Also, children who live with parents who smoke have easier access to cigarettes than children who live with non-smoking parents. If smoking is also associated with divorce, parameter estimates may be biased.* [5, p. 61]

This information can be formalized as follows:

37	$PS \neq SP$	–	PREM	$\emptyset$
38	$SP \neq SI$	–	PREM	$\emptyset$
39	$\neg(PS \amalg SP)$	–	PREM	$\emptyset$
40	$\neg(SP \amalg SI)$	–	PREM	$\emptyset$

A parallel way of thinking about the role of the smoking status of parents has been leading to a third analogous research question, namely $?[(PS \amalg SI \mid SP)]$. And again this proposal has been withdrawn on the basis of the statistical results.

A last possibility was that those three variables would act all together as a screener off between parental separation and smoking initiation. Maybe they are as a triple common cause able to explain the association between parental separation and smoking initiation? This supposition leads to the research question $?[(PS \amalg SI \mid (HI \wedge PE \wedge SP)]$.

The statistical results nevertheless show that the association between parental separation and smoking initiation stays present, even when controlling on the three variables together. Nonetheless, the possible indirect effect of those variables was a reason to integrate all of them as control variables when investigating the causal path between parental separation and smoking initiation.

3.3. *Searching for Mediating Variables*

3.3.1. *Hypotheses and Research Questions*

Since no common causes that could explain away the association between parental separation and smoking initiation were found, the point of departure for the second phase of the study is still the initial finding (cf section 3.2) that there is a causal path from parental separation to smoking initiation:

1	$\neg(PS = SI)$	–	PREM	$\emptyset$
2	$\neg(PS \amalg SI)$	–	PREM	$\emptyset$
3	$(PS \rightarrow SI) \vee (SI \rightarrow PS) \vee (PS = SI)$	2	Rc	$\mathcal{SO}_{PS\,SI}$
4	$(PS \rightarrow SI) \vee (SI \rightarrow PS)$	1,3	CL	$\mathcal{SO}_{PS\,SI}$
5	$\neg(SI \rightarrow PS)$	–	BK	$\emptyset$
6	$PS \rightarrow SI$	4,5	CL	$\mathcal{SO}_{PS\,SI}$

In the further search, Kirby considers 6 possible mediating variables: DP (depressed mood), R (rebelliousness) and SE (self esteem) as the possible mediating variables pertaining to psychological distress; PCC (parent child closeness), PSV (parental supervision) and PI (peer influence) as the possible mediating variables pertaining to socialization.

Why these six variables? Kirby relies on the one hand on the risk factors for smoking initiation as identified by previous research and on the other hand on socialization theory and stress theory. According to socialization theory, the primary reason why children who experience the divorce of their parents are, on average, worse off than other children is that the family types that result from parental separation are less effective in providing a secure and consistent environment for socialization. Stress theory on the other hand, suggests that children are generally better off in two-parent families because children in two-parent families do not experience the conflict and instability resulting from the separation of their parents. Kirby extends the ideas of socialization theory and stress theory to develop hypotheses regarding the possible effect of parental separation on smoking initiation. He considers the six above mentioned risk factors for smoking initiation as potential mediating variables. He starts from two hypotheses. A first hypothesis draws on stress theory:

> *Parental separation increases the likelihood that adolescents will start smoking by increasing depressed mood and rebelliousness and decreasing self-esteem.* [5, p. 59]

This hypothesis implies the consideration of the following causal relations:

A $?[(PS \rightarrow DM) \wedge (DM \rightarrow SI)]$
B $?[(PS \rightarrow R) \wedge (R \rightarrow SI)]$
C $?[(PS \rightarrow SE) \wedge (SE \rightarrow SI)]$

Another research question which follows from this is question D, namely if depressed mood screens off the correlation between parental separation and smoking initiation:

D $?[(PS \amalg SI|DM)]$

Only the combination of a positive answer to both questions A and D result in the rejection of a direct causal influence of parental separation on smoking initiation.

An analogous way of reasoning can be made towards rebelliousness and self esteem. As a consequence, parallel research questions can be put forward:

E $?[(PS \amalg SI|R)]$
F $?[(PS \amalg SI|SE)]$

A second additional hypothesis considered by Kirby, draws on stress theory:

> *Parental separation increases the likelihood that adolescents will start smoking by decreasing closeness to parents, decreasing parental supervision, and increasing adolescents' association with friends who smoke.* [5, p. 59]

In other words, Kirby is pondering the following causal relations:

G $?[(PS \rightarrow PCC) \wedge (PCC \rightarrow SI)]$
H $?[(PS \rightarrow PSV) \wedge (PSV \rightarrow SI)]$
I $?[(PS \rightarrow PI) \wedge (PI \rightarrow SI)]$

Those questions lead again to the following three questions. The respective combinations of possible positive answers to the previous questions with possible positive answers to the following questions, can assure us about the absence of a direct causal influence of parental separation on smoking initiation:

J $?[(PS \amalg SI|PCC)]$
K $?[(PS \amalg SI|PSV)]$
L $?[(PS \amalg SI|PI)]$

Once we know the answers to these questions, and if a screener off between parental separation and smoking initiation has not yet been found, another possibility remains. The variables that proved to be mediating, can act all together as a screener off between parental separation and smoking initiation in adolescents. By consequence, also this possibility will have to be examined.

Before going on, it can be useful to know a little more about the conclusions of the study. Those have been summarized by Kirby as follows.

> *Three variables emerge as significant mediators between parental separation and smoking initiation. First, parental separation is associated with an increase in the number of friends adolescents have who smoke, and this in turn increases the likelihood that an adolescent will start smoking. Parental separation also elevates both the level of depressed mood and the level of rebelliousness in adolescents, and these things in turn increase the likelihood of smoking initiation. Though a significant amount of the total effect of parental separation on smoking initiation is explained by the mediating variables in my model, a majority of the effect is direct.* [5, p. 67]

3.3.2. *Analysis of Statistical Data*

With the above research questions in the back of his mind, Kirby analysed the statistical data. For matters of space, we will abbreviate the first 28 lines of the sequel of the corresponding proof to $\neg(X = Y)$, stating that none of the variables is equivalent to one of the others.

7-34	$\neg(X = Y)$	–	PREM	$\emptyset$
35	$\neg(PS \amalg DM)$	–	PREM	$\emptyset$
36	$\neg(PS \amalg R)$	–	PREM	$\emptyset$
37	$\neg(PS \amalg SE)$	–	PREM	$\emptyset$
38	$\neg(R \amalg SI)$	–	PREM	$\emptyset$
39	$\neg(DM \amalg SI)$	–	PREM	$\emptyset$
40	$(SE \amalg SI)$	–	PREM	$\emptyset$

The finding that

> *all three measures of psychological distress are significantly affected by parental separation.* [5, p. 66]

corresponds with premises 35 to 37, whereas premises 38 to 40 are justified by the experience that

> *Not surprisingly, an increase in rebelliousness...is associated with an increase in the likelihood that adolescents will start smoking. Similarly, an increase in depressed mood is associated with an increase in the likelihood that an adolescent will start smoking. Self-esteem however, shows no association with smoking initiation.* [5, p. 66]

The lack of a correlation between self esteem and smoking initiation (cf line 40) points to the impossibility of the derivation of a causal path between divorce and smoking initiation that goes via self esteem. This implies there will anyhow be no way to derive hypothesis C.

Lines 41 till 46 are introduced for the sake of completeness because depressed mood, rebeliousness and self esteem are examined as three manifestations of psychological distress. As such, they are dependent. But conditional on parental separation, those three manifestations seem to manifest themselves in diffusing ways. Above all, with those premises we explicate we are not in search for causal relations among these three measures:

41	$(DM \amalg R \mid PS)$	–	PREM	$\emptyset$
42	$(DM \amalg SE \mid PS)$	–	PREM	$\emptyset$
43	$(R \amalg SE \mid PS)$	–	PREM	$\emptyset$
44	$\neg(DM \amalg R)$	–	PREM	$\emptyset$
45	$\neg(DM \amalg SE)$	–	PREM	$\emptyset$
46	$\neg(R \amalg SE)$	–	PREM	$\emptyset$

The premises in accordance with the second hypothesis of Kirby, and thus concerning the possible mediating variables parent-child closeness, parental supervision and peer influence, can be formalized as follows:

47	$\neg(PS \amalg PCC)$	–	PREM	$\emptyset$
48	$\neg(PS \amalg PI)$	–	PREM	$\emptyset$
49	$(PS \amalg PSV)$	–	PREM	$\emptyset$
50	$\neg(PI \amalg SI)$	–	PREM	$\emptyset$
51	$\neg(PSV \amalg SI)$	–	PREM	$\emptyset$
52	$(PCC \amalg SI)$	–	PREM	$\emptyset$

Line 47 till 49 are justified by the following findings:

> *Parental separation has a significant direct effect on two of the mediating variables pertaining to socialization. Adolescents who experience the divorce or separation of their parents...report feeling significantly less close to their fathers... Parental separation also seems to influence the number of friends adolescents have who smoke. Adolescents report an increase in the number of friends they have who smoke. Parental separation does not, however, have a significant effect on the extent to which adolescents feel close to their mothers. Furthermore, there is no evidence that parental separation affects the amount of time that an adolescent spends at home without the presence of a mother.* [5, p. 64]

Line 50 till 52 are based on the following facts:

> *two of the mediating variables pertaining to socialization are significant predictors of smoking initiation. Increases in the amount of time adolescents spend unsupervised are positively related tot he likelihood that they will start smoking. Also, an increase in the number of friends adolescents have that smoked is strongly and positively related to the smoking initiation. However neither change in mother-child closeness nor change in father-child closeness emerges as a significant predictor of smoking initiation in adolescents.* [5, p. 64]

Because no correlation, and by consequence no causal relation can be found here between on the one side parental separation and parental supervision (cf line 49), and on the other side parent-child closeness and smoking initiation (cf line 52), research questions I and J can be dropped. The only mediating variable pertaining to socialization theory that will stay upright is peer influence.

For the sake of completeness another six premises are introduced because PCC, PSV and PI are three possible measures of the sociological impact, but conditional on parental separation those three variables seem to manifest themselves in diffusing ways. And again, with those premises we explicate we are not in search for causal relations among these three measures.

53	$(PCC \amalg PI \mid PS)$	–	PREM	$\emptyset$
54	$(PCC \amalg PSV \mid PS)$	–	PREM	$\emptyset$
55	$(PSV \amalg PI \mid PS)$	–	PREM	$\emptyset$
56	$\neg(PCC \amalg PI)$	–	PREM	$\emptyset$
57	$\neg(PCC \amalg PSV)$	–	PREM	$\emptyset$
58	$\neg(PSV \amalg PI)$	–	PREM	$\emptyset$

In accordance with the rules of ALcd, the following derivations can be made according to the three remaining possible mediating variables depressed mood, rebelliousness and peer influence:

59	$(PS \rightarrow DM) \vee (DM \rightarrow PS)$ $\vee (PS = DM)$	35	Rc	$\mathcal{SO}_{PS\,DM}$
60	$(PS \rightarrow DM) \vee (DM \rightarrow PS)$	7-34,59	CL	$\mathcal{SO}_{PS\,DM}$
61	$\neg(DM \rightarrow PS)$	BK	$\emptyset$	
62	$PS \rightarrow DM$	60,61	CL	$\mathcal{SO}_{PS\,DM}$
63	$(PS \rightarrow R) \vee (R \rightarrow PS) \vee (PS = R)$	36	Rc	$\mathcal{SO}_{PS\,R}$
64	$(PS \rightarrow R) \vee (R \rightarrow PS)$	7-34,63	CL	$\mathcal{SO}_{PS\,R}$
65	$\neg(R \rightarrow PS)$	BK	$\emptyset$	
66	$PS \rightarrow R$	64, 65	CL	$\mathcal{SO}_{PS\,R}$
67	$(PS \rightarrow PI) \vee (PI \rightarrow PS) \vee (PS = PI)$	48	Rc	$\mathcal{SO}_{PS\,PI}$
68	$PS \rightarrow PI \vee PI \rightarrow PS$	7-34,67	CL	$\mathcal{SO}_{PS\,PI}$
69	$\neg(PI \rightarrow PS)$	BK	$\emptyset$	
70	$PS \rightarrow PI$	68,69	CL	$\mathcal{SO}_{PS\,PI}$

In line 61, 65, and 69 some background knowledge has been introduced that is presupposed in this study to make further derivations. In this case, we were looking for causal pathways between parental separation and smoking initiation. By consequence, parental separation can only be a cause of other variables here, and not an effect variable itself. That is what justifies those lines.

At this point, the proposed causal paths between parental separation and possible mediating variables are analysed. The following things concerning the paths from those variables to smoking initiation can be derived:

75	$(R \rightarrow SI) \vee (SI \rightarrow R) \vee (R = SI)$	38	Rc $\mathcal{SO}_{R\,SI}$
76	$(R \rightarrow SI) \vee (SI \rightarrow R)$	7-34,75	CL $\mathcal{SO}_{R\,SI}$
77	$(DM \rightarrow SI) \vee (SI \rightarrow DM)$ $\vee (DM = SI)$	39	Rc $\mathcal{SO}_{DM\,SI}$
78	$(DM \rightarrow SI) \vee (SI \rightarrow DM)$	7-34,77	CL $\mathcal{SO}_{DM\,SI}$
79	$R \rightarrow SI$	7-34,41,76,78	DR2 $\mathcal{SO}_{R\,SI}$
80	$DM \rightarrow SI$	7-34,41,76,78	DR2 $\mathcal{SO}_{DM\,SI}$
81	$(PI \rightarrow SI) \vee (SI \rightarrow PI) \vee (PI = SI)$	50	Rc $\mathcal{SO}_{PI\,SI}$
82	$(PI \rightarrow SI) \vee (SI \rightarrow PI)$	7-34,79	CL $\mathcal{SO}_{PI\,SI}$

83	$(PSV \rightarrow SI) \vee (SI \rightarrow PSV)$ $\vee (PSV = SI)$	51	Rc $\mathcal{SO}_{PSV\,SI}$
84	$(PSV \rightarrow SI) \vee (SI \rightarrow PSV)$	7-34,83	CL $\mathcal{SO}_{PSV\,SI}$
85	$PI \rightarrow SI$	1,55,82,84	DR2 $\mathcal{SO}_{PI\,SI}$
86	$PSV \rightarrow SI$	1,55,82,84	DR2 $\mathcal{SO}_{PSV\,SI}$

It is clear that we need the causal path between parental supervision and smoking initiation only to derive the path between peer influence and smoking initiation. It is of no further use, because no correlation between parental separation and parental supervision has been found.

Looking back to the research questions, we can see that research questions A, B and I can be affirmed by the conjunction of respectively line 62 and 80, line 66 and 79, and line 70 and 85:

87	$(PS \rightarrow DM) \wedge (DM \rightarrow SI)$	62,80	CL	$\mathcal{SO}_{PS\,DM} \cup \mathcal{SO}_{DM\,SI}$
88	$(PS \rightarrow R) \wedge (R \rightarrow SI)$	66,79	CL	$\mathcal{SO}_{PS\,R} \cup \mathcal{SO}_{R\,SI}$
89	$(PS \rightarrow PI) \wedge (PI \rightarrow SI)$	70,85	CL	$\mathcal{SO}_{PS\,PI} \cup \mathcal{SO}_{PI\,SI}$

Furthermore, nowhere in the data is there any evidence that one or another variable or a combination thereof explains away the correlation between parental divorce and smoking initiation. By consequence, although part of the influence of parental separation on smoking initiation is mediated by depressed mood, rebelliousness and peer influence, part of the influence remains direct (cf line 6). All this gives us the full picture, in accordance with the conclusions of Kirby as cited above.

4. *Concluding Remarks and the Incorporation of Mechanisms*

We succeeded in demonstrating the applicability of the logic ALcd to the formalization of the research process of Kirby. We are convinced of the relevance of his study as an example of the way a lot of researchers go on in their work. This study is surely not an isolated case. A lot of other research in the field of (social) science will be able to be formalized in an analogous way. By consequence, we are convinced we succeeded in affirming the usefulness of this logic in formalizing the way in which humans think from association to causation, in accordance with our presuppositions.

Nonetheless, we like to make some comments to our work. First, it would be much more elegant to specify the rule QRc. As we proposed it now, we are able to introduce a question at any moment in the proof. Implicitly

however, we already derived the questions at some specific moments, given the presence of some significant lines. We can convey this implicitly used rules as follows [2]:

QRc1 $\alpha \rightarrow \beta$,
$\gamma \rightarrow \alpha$,
$\gamma \rightarrow \beta$
$\overline{?[(\alpha \amalg \beta|\gamma)]}$

QRc2 $\alpha \rightarrow \gamma$,
$\gamma \rightarrow \beta$,
$\alpha \rightarrow \beta$
$\overline{?[(\alpha \amalg \beta|\gamma)]}$

The answer to the question derived by QRc1, makes it possible to distinguish between the two possible situations of figure 1 and 2.

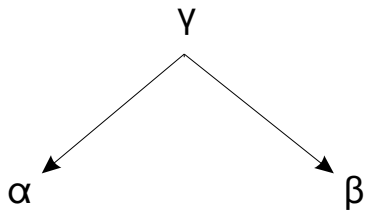

Figure 1. Possibility 1 in case of QRc1.

Knowing the answer to the question derived by QRc2, we can decide between the situations of figure 3 and 4 :

We could also formulate similar rules for combinations of variables as possible screeners off. It is clear they will be analogous, we will not write them out here. Another option for the formalization of the heuristic process would be to develop goal directed proofs for the logic ALcd, analogous to the proofs for ACluN1 as developed in [2].

However, not all surveys will be so easy to transform in an ALcd-reasoning process. An important reason for this is the use of mechanisms in scientific explanations. For this, it would be worth making some adaptations to this causal discovery logic such that it is able to take the role of mechanisms into

[2] The idea for the introduction of this kind of rules has been derived from the adaptive logic for question evocation from [6].

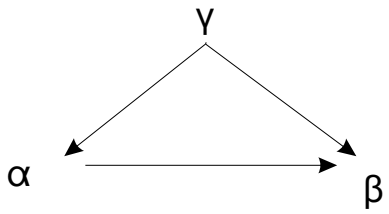

Figure 2. Possibility 2 in case of QRc1.

Figure 3. Possibility 1 in case of QRc2.

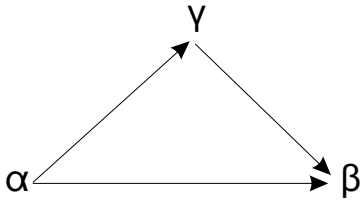

Figure 4. Possibility 2 in case of QRc2.

account. Extending ALcd with some rules that make it possible to incorporate mechanisms, would supply us with a stronger instrument when dealing with social science. By using the word 'mechanisms' we intend here to refer to those intervening processes between a cause and an effect which can't be captured in correlational terms. In other words, we are not talking about intervening variables that just screen off a cause and effect as within the ALcd-reasoning context. We talk about some non-empirical evidence that explains the connection between the cause and the effect. As has been put forward by Schutt:

> *Confidence in a conclusion that there is a causal connection between two variables will be strengthened if a mechanism — some discernible means of creating a connection — can be identified. Such mechanisms ... help us to understand how variation in the independent variable(s) results in variation in the dependent variable.* [7, p. 121]

This quote already points to a first way in which mechanisms are helpful. When we are convinced of having found a mechanism that is evidence for the causal connection between two variables, our belief in this causal connection will be strengthened. This means in practice there will be no reasons left for a further search for possible screeners-off. In the concrete, finding a mechanism gives reasons to remove the fifth element of the line in the proof, namely the condition $\mathsf{SO}_{\alpha,\beta}$. For example, Christiaens [3] found a correlation between the use of medicines in case of ADHD, and the child experiencing ADHD as a problem. At first sight, we would expect the reverse. The use of medicines should cause to lessen the experience of ADHD as a problem. Nonetheless, Christiaens also found a mechanism clarifying the correlation, namely the self-categorization theory. As long as the children are not forced to recognize they are different from others (and this is exactly what *does* happen when they are obliged to use a medicine to change their behaviour), the children with ADHD don't experience themselves as different from other children, and by consequence do not see themselves as "problematic" children. When they nonetheless take medicines, they know they need to use it because of "their problem". The medicine will then become a justification of their conduct. As this theory is for Christiaens convincing evidence for the causal relation between medicalizing ADHD and experiencing it as a problem, she will not search for further screeners-off between those two factors.

Another way in which we can incorporate the information of mechanisms is in identifying spurious or coincidental correlations. When a causal connection between two variables is supposed, but no plausible mechanism at all can be found that elucidates this connection, we need to revise the supposition of a direct relation and search for mediating variables, common causes, or some other means in which the variables influence indirectly each other. This means we can add a rule which states that a line that proposes a causal connection, for which there can nevertheless not a single clarifying mechanism be found, has to be marked. In that case, only further information about correlations and screeners off can help us out. For example, when we find a possible causal connection between an increase in churches and an increase in prostitution, people would have difficulty believing in a causal relation between those factors. They will automatically search for more information that can clarify this correlation. In our proof, the line that states

a direct causal connection between the increase in churches and the increase in prostitution, will be marked. Further information can then for example demonstrate that both factors are an effect of the common cause 'increase in population'.[3]

Centre for Logic and Philosophy of Science
Ghent University, Belgium
E-mail: Leen.DeVreese@rug.ac.be,
Erik.Weber@rug.ac.be

REFERENCES

[1] W. Ahn and C. W. Kalish e.a. The role of covariation versus mechanism information in causal attribution. *Cognition*, 54:299–352, 1995.

[2] Diderik Batens. Some computational aspects of inconsistency-adaptive logics. CLE e-Prints, 2(7):15, p. 2002, http://www.cle.unicamp.br/e-prints/abstract_16.html.

[3] Wendy Christiaens. Identiteit en de betekenis van overbeweeglijkheid. *Tijdschrift voor sociologie*, 22 (2):137–151, 2001.

[4] Maarten Van Dyck. Causal discovery using adaptive logics. *Logique et Analyse*, this volume, p. xxx, 2002.

[5] James B. Kirby. The influence of parental separation on smoking initiation in adolescents. *Journal of Health and Social Behavior*, 43:56–71, 2002.

[6] Joke Meheus. Adaptive logics for question evocation. *Logique et Analyse*, 173–174–175: 135–164, 2001 (Appeared 2003).

[7] Russel K. Schutt, editor. *Investigating the Social World. The Process and Practice of Research*. The Pine Forge Press series in research methods and statistics. Pine Forge Press, Thousand Oaks (Calif.), 1996.

[3] Example in [1].