

김세화 교수의 반론과 활성화 경로 이론의 제한 및 확장*

김 성 수

【국문요약】 인과 관계가 결과와 원인 간의 반사실적 의존 관계로 분석될 수 있다는 직관은 설득력이 있다. 활성화 경로 이론은 단순한 형태의 반사실적 인과 이론의 문제점을 피하면서도 이 직관을 유지하기 위해 제안된 이론이다. 하지만 이 이론 역시 심각한 반례에 직면한다. 김세화 교수는 최근의 논문에서 이러한 반례를 반박하고자 제시된 기존의 해결책이 갖는 문제점을 설득력 있게 비판하였다. 이 논문은 김세화 교수의 비판을 논의하고 더 나아가 활성화 경로 이론에 대한 반례를 극복할 수 있는 또 다른 해결책을 논의한다. 특히 활성화 경로 이론의 적용 범위를 제한하고 이렇게 제한된 이론을 다시 확장하는 방식을 제시함으로써 활성화 경로 이론을 그 반례로부터 방어하는 방법을 제안한다.

【주요어】 인과, 반사실적 인과론, 활성화 경로, 반례, 김세화

접수일자: 2015.04.18 심사 및 수정완료일: 2015.05.20 게재확정일: 2015.06.01

* 이 논문의 심사를 맡아 유익한 비판과 도움말을 해주신 익명의 심사위원님
들께 감사드립니다.

1. 들어가는 말

김세화 교수는 최근 논문 “김성수 교수의 활성화 경로 이론에 대한 변호와 그에 대한 반론”에서¹⁾ 활성화 경로 이론에 대한 기존의 반례에 대응하여 김성수(2012)가 제시한 해법을 설득력 있게 비판하였다. 이 비판이 보여주는 것은, 활성화 경로 이론은 그 이론의 반례를 극복할 수 있는 또 다른 방법이 제시되지 않는 한 심각한 문제가 있는 이론이라는 점이다. 이 논문의 목적은 김세화 교수의 비판을 논의하고 더 나아가 활성화 경로 이론을 옹호할 수 있는 또 다른 가능성이 있는지를 살펴보는 것이다. 이를 위해 2절에서는 먼저 활성화 경로 이론과 그 반례들을 간략히 정리한다. 3절은 기존의 해법에 대한 김세화 교수의 비판을 논의한다. 좀 더 자세히 말하면, 이 절에서는 홀의 사례와 관련된 비판이 기존의 해결책 전반에 문제가 있음으로 보이는 결정적인 비판임에 동의한다. 반면 치명적 해독제의 사례와 관련된 비판은 논쟁의 여지가 있음을 제시하고, 그럼에도 불구하고 이 사례에 대한 기존의 해결책은 여전히 문제가 있음을 지적할 것이다. 4절은 활성화 경로 이론의 문제점을 극복하는 방법을 모색한다. 이 시도는 두 단계로 이루어진다. 첫 번째는 활성화 경로 이론의 적용 범위를 ‘인접’ 인과 관계로 제한하는 것이다. 두 번째는 인접 인과 관계들로 구성된 인과 사슬 혹은 인과 망(network)에서 성립하는 ‘경유’ 인과 관계를 규명하는 방법을 제시하는 것이다. 그리고 이렇게 제한되고 확장된 형태의 이론이 기존의 활성화 경로 이론에 대한 반례를 해결할 수 있는지 살펴볼 것이다.

1) 김세화 (2015).

2. 활성화 경로 이론과 반례들

원인과 결과 간에 반사실적 의존 관계가 성립한다는 생각은 직관적으로 설득력이 있다. 하지만 인과 관계를 단순한 반사실적 의존 관계로 분석할 수 없다는 것은 이미 잘 알려진 사실이다. 예를 들어 다음의 경우를 생각해보자.

(선점) 철수와 영희 둘 다 자신이 든 돌을 유리창에 던지려고 했지만 철수가 먼저 돌을 던지고 이를 본 영희는 돌을 던지지 않기로 한다. 하지만 만약 철수가 돌을 던지지 않았더라면 영희가 돌을 던져 유리창이 깨졌을 것이다. 한편 철수가 던진 돌은 유리창을 향해 날아가고 유리창이 깨진다.²⁾

(선점)에서 유리창이 깨진 것의 원인은 철수가 돌을 던진 사건이지만, 이러한 실제 원인이 발생하지 않을 경우 영희가 돌을 던졌을 것이기 때문에 유리창은 여전히 깨졌을 것이다. 이렇듯 (선점)에서는 실제 원인이 발생하지 않은 경우에도 동일한 결과가 발생하도록 하는 요인이 존재하므로 원인과 결과 간에 반사실적 의존 관계가 성립하지 않는다.

그런데 원인과 결과 간의 반사실적 의존 관계를 방해하는 요인을 일정한 방식으로 컨트롤하면, 원인과 결과 간에 의존 관계가 성립할 수 있다. 예를 들어, (선점)에서 영희가 실제로는 돌을 던지지 않았다는 사실을 고정한 상태에서 (즉 이 사실이 성립하는 그런 가능 세계에서) 철수가 돌을 던지지 않았더라면 유리창은 깨지지 않았을 것이다. 활성화 경로 이론은 ‘인과 모델’을 사용하는 인과 구조식 접근법의 일종으로, 특히 ‘일정한 방식의 컨트롤’ 하에서 성립하는 반사실적 의존 관계를 통해 인과 관계를 파악하고자 하는 이론

2) 김성수 (2012), pp. 262-263.

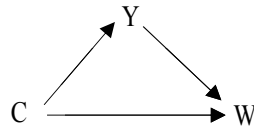
이다.

(선점)에서 중요한 요인 혹은 ‘변수’는 철수가 돌을 던지는지의 여부(C), 영희가 돌을 던지는지의 여부(Y) 그리고 유리창이 깨지는지의 여부(W)이다. 그리고 이들 변수 간에는 $Y = \sim C$ 와 (즉 철수가 돌을 던지지 않는 경우 영희는 돌을 던지고, 반대로 철수가 돌을 던지는 경우 영희는 돌을 던지지 않는 관계와) $W = C \vee Y$ (즉 철수가 돌을 던지거나 혹은 영희가 돌을 던지는 경우 유리창이 깨진다는) 관계가 성립하는데, 이 관계를 ‘ $Y = \sim C$ ’와 ‘ $W = C \vee Y$ ’의 구조식이 보여준다. 그리고 이 사례에서 변수 C는 다른 변수에 의존하지 않는데 그런 경우에는 그 변수가 실제로 갖는 값을 보여줌으로써($C = 1$) C에 대한 구조식을 나타내면 된다. 이제 변수 {C, Y, W}와 이들 간에 성립하는 위의 구조식들을 이 사례에 대한 인과 모델이라고 하자. 이 인과모델과 그것을 시각적으로 보여주는 그래프는 다음과 같다.

$$C = 1$$

$$Y = \sim C$$

$$W = C \vee Y$$



이와 같은 인과 모델이 주어지면, 활성화 경로 이론은 인과 관계를 다음과 같이 분석한다.

활성 경로 이론: $X = x$ 는, X에서 시작하여 Y로 이어지는 활성화 경로가 적어도 하나 존재하는 경우 그리고 단지 그 경우에만 $Y = y$ 의 원인이다.

여기서 ‘활성 경로’는 다음과 같이 정의된다.

활성 경로: X에서 시작하여 Y로 연결되는 경로는, 그 경로에 속하지 않은 변수의 값을 그 변수가 주어진 상황에서 실제로 갖는 값으로 고정한 상태에서, Y의 값이 X의 값에 반사실적으로 의존하는 경우 그리고 단지 그 경우에만 활성화 경로이다.

위 그림에서 C에서 W로 이어지는 경로는 C-Y-W와 C-W 둘이다. 이제 $C = 1$ (철수가 돌을 던진 것)이 $W = 1$ (유리창이 깨진 것)의 원인인지의 여부는 이 두 경로 중 적어도 한 경로가 활성화 경로인지의 여부에 의해 결정된다. C-Y-W 경로의 경우 그 경로에 속하지 않은 변수가 없으므로, 어떤 변수의 값도 고정하지 않은 상태에서 C와 W 간의 반사실적 의존 관계를 검토해야 한다. 그런데 $C = 1$ 을 $C = 0$ 으로 변화시키면 Y의 값은 0에서 1로 변하게 되므로 W의 값은 아무 변화 없이 $W = 1$ 에 머문다. 따라서 이 경로는 활성화 경로가 아니다. 그리고 이 경로는 기존의 단순한 형태의 반사실적 인과론이 살펴볼 것을 요구하는 (유일한) ‘경로’이며 단순한 형태의 이론에 대한 반례를 제공하는 ‘경로’이다. 반면 C-W 경로는 활성화 경로이다. 이 경로에 속하지 않은 Y를 그것의 실제 값인 $Y = 0$ (영희가 돌을 던지지 않음)에 고정한 상태에서 X를 $X = 1$ 에서 $X = 0$ 으로 변화시키면 W는 $W = 1$ 에서 $W = 0$ 으로 변한다. 따라서 활성화 경로 이론에 따르면, $C = 1$ 은 $W = 1$ 의 원인이며 이는 직관과 부합하는 답이다.

이렇듯 활성화 경로 이론은 인과 관계를 기본적으로 반사실적 의존 관계로 분석하는 한편 그 과정에서 생길 수 있는 문제를 피해갈 수 있는 방법을 제공하고자 한다. 하지만 이 이론 역시 반례를 갖는다.

(치명적 해독제) 암살자가 보스의 커피에 독을 넣었고 이를 본 보디가드는 커피에 해독제를 넣었다. 만약 보디가드가 해독제를 넣

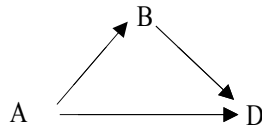
지 않았더라면 보스는 커피를 마시고 죽었을 것이다. 반면 해독제는 그것만 마셨을 경우 치명적이며 커피에 독이 있지 않았더라면 해독제는 보스를 죽였을 것이다. 독약과 해독제는 그것만 마시면 치명적이지만 함께 마시면 서로를 중화한다. 보스는 커피를 마셨고 살아남았다.³⁾

변수 A는 암살자가 커피에 독약을 넣는 경우 1 (아닌 경우 0), B는 보디가드가 커피에 해독제를 넣는 경우 1 (아닌 경우 0), D는 보스가 죽는 경우 1 (아닌 경우 0)이라고 할 때, (치명적 해독제)의 인과 구조식 및 관련된 그래프는 다음과 같다.

$$A = 1$$

$$B = A$$

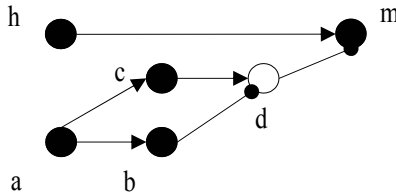
$$D = (A \ \& \ \sim B) \ \vee \ (\sim A \ \& \ B)$$



B를 실제 값인 $B = 1$ 에 고정했을 때, A의 값을 A의 실제 값인 $A = 1$ 에서 $A = 0$ 으로 변화시키면 D의 값은 그것의 실제 값인 $D = 0$ 에서 $D = 1$ 로 바뀐다. 따라서 활성 경로 이론에 따르면 암살자가 커피에 독약을 넣은 것($A = 1$)은 보스가 살아남은 것($D = 0$)의 원인이 되어버리지만, 이는 직관적으로 잘못된 답으로 보인다.

이에 더해 활성 경로 이론은 다음과 같은 반례에도 직면한다.

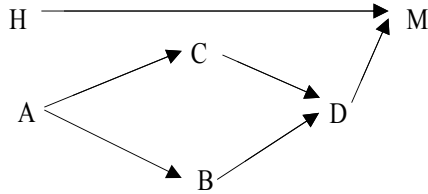
(홀의 사례)



³⁾ Menzies (2004), p. 825. 김성수 (2012), pp. 274-275에서 재인용.

홀이 제시한 이 뉴런 다이어그램에서⁴⁾, a-뉴런이 켜진 것은 b-뉴런과 c-뉴런을 켜지게 한다. 그리고 c-뉴런은 d-뉴런이 켜지도록 촉진하는 반면 b-뉴런은 d-뉴런이 켜지지 않도록 억제하며, 그 결과 d-뉴런은 켜지지 않는다. 이 사례에서 a-뉴런이 켜진 것은 d-뉴런이 켜지는 것을 촉진하는 동시에 억제하는 역할을 하므로 a-뉴런이 켜진 것은 d-뉴런이 꺼진 것의 원인이 아니다. 이와 유사하게, m-뉴런이 켜지는 것은 d-뉴런이 켜져 m-뉴런을 억제하는지 여부에 달려 있으므로 a-뉴런이 켜진 것은 m-뉴런이 켜진 것의 원인이 아니다. 하지만 활성화 경로 이론은 잘못된 답을 내놓는다.

H = 1
 A = 1
 B = A
 C = A
 D = C & ~B
 M = H & ~D



C를 실제 값인 1에 고정한 상태에서 A의 값을 $A = 0$ 으로 변화시키면 $B = 0$ 이 되어 $D = 1$ 로 변화된다. 따라서 A-B-D는 활성화 경로이다. 이와 유사하게, C를 실제 값인 1에 고정한 상태에서 A의 값을 $A = 0$ 으로 변화시키면 $D = 1$ 으로 변하고 따라서 M의 값 역시 $M = 0$ 으로 변하므로 A-B-D-M 역시 활성화 경로이다. 활성화 경로 이론에 따르면 a-뉴런이 켜진 것은 d-뉴런이 꺼진 것의 원인이나 m-뉴런이 켜진 것의 원인이지만, 이는 우리의 직관에 반하는 잘못된 답이다.

⁴⁾ Hall (2007), p. 120. 김성수 (2012), p. 279에서 재인용.

3. 반례에 대한 해결책과 김세화 교수의 비판

김성수(2012)는 (홀의 사례)가 실은 진정한 반례가 될 수 없다고 주장하기 위해 이와 동일한 구조를 갖는 사례를 살펴볼 것을 제안 하는데, 그 사례는 다음과 같이 정리될 수 있다.

(독이 이미 들어 있는 커피에 넣은 촉매제) 암살자가 커피에 독을 넣는다. 이를 본 조력자는 이미 독이 들어 있는 커피에 촉매제를 넣고, 이(암살자가 커피에 독을 넣는 것)를 본 보디가드는 커피에 해독제를 넣는다. 커피에 넣은 독이 효력을 발휘하기 위해서는 촉매제가 필요하지만 해독제는 그렇지 않다. 해독제와 촉매제는 그것만 마시든 함께 마시든 무해하다. 만약 보디가드가 해독제를 넣지 않았더라면 보스는 커피를 마시고 죽었을 것이다. 실제로 보스는 커피를 마셨고 살아남았다. 덕분에 보스는 해설자의 도움으로 오전 미팅을 성공적으로 수행하였다.⁵⁾

(홀의 사례)에서 변수 D의 구조식($D = C \ \& \ \sim B$)의 값은 변수 A의 값과 상관없이 단지 변수 C와 B의 값에 의해서만 결정된다. 위의 사례가 홀의 사례에 대응하기 위해서는 변수 A는 암살자가 커피에 독을 넣는지의 여부, 변수 B는 보디가드가 커피에 해독제를 넣는지의 여부, 변수 C는 이미 독이 들어 있는 커피에 조력자가 촉매제를 넣는지의 여부 그리고 변수 D는 보스가 죽는지 여부로 보면 된다.

(홀의 사례)가 활성화 경로 이론에 대한 반례인 이유는 C의 값을 실제 값인 $C = 1$ (이미 독이 들어 있는 커피에 조력자가 촉매제를 넣은 것)에 고정할 때, A의 값을 $A = 0$ (암살자가 독을 넣지 않음)으로 변화시키면 D의 값 역시 $D = 1$ (보스가 죽음)으로 변하고 따라서 암살자가 독을 넣은 것이 보스가 살아남은 것의 원인이 되어

⁵⁾ 김세화 (2015), p. 143.

버리기 때문이었다. 하지만 *이미 독이 들어 있는* 커피에 조력자가 촉매제를 넣었음을 고정된 상태에서 암살자가 커피에 독을 넣지 않는 것을 상정하는 것은 실제와 아주 동떨어진 상황이다. 그리고 (홀의 사례)가 실제 상황과는 아주 동떨어진 상황에 대해 생각하기를 요구한다는 것은 이 사례가 설득력 있는 반례가 될 수 없다는 것을 뜻한다.

그러나 김세화 교수는 이 해결책은 “홀의 다이어그램 자체에 대한 반박이라기보다는 이 다이어그램을 예시하는 한 구체적인 사례에 대한 반박에 그칠 수 있다”⁶⁾는 점을 지적한다. 다음의 사례를 생각해보자.

(명령하는 암살자) 암살자가 조력자에게 보스의 커피에 독을 넣으라고 명령한다. 이 명령을 듣고 조력자는 커피에 독을 넣는다. 한편 암살자가 조력자에게 명령하는 것을 본 보디가드는 커피에 해독제를 넣는다. 해독제는 그것만 마시는 경우 무해하다. 실제로 보스는 커피를 마셨고 살아남았다. 덕분에 보스는 해설자의 도움으로 오전 미팅을 성공적으로 수행하였다.⁷⁾

(홀의 사례)에서 A는 암살자가 조력자에게 명령을 내리는지의 여부, B는 보디가드가 해독제를 넣는지의 여부, C는 조력자가 독을 넣는지의 여부 그리고 D는 보스가 죽는지의 여부라고 하자. 이제 C를 실제 값인 $C = 1$ (조력자가 커피에 독을 넣은 것)에 고정하면, D는 A에 반사실적으로 의존한다. 따라서 활성화 경로 이론에 따르면 암살자가 명령을 내린 것은 보스가 죽지 않은 것의 원인이라는 반직관적인 답이 도출된다. 하지만 (독이 이미 들어 있는 커피에 넣은 촉매제)와는 달리, 이 경우 C의 값을 고정된 상태에서 A의 값을 변화시키는 것은 실제 상황과 아주 동떨어진 상황이 아니다. 김

6) 김세화 (2015), p. 148.

7) 김세화 (2015), p. 148.

세화 교수가 지적하듯 “암살자가 명령을 내리지도 않았는데 조력자가 커피에 독을 넣는 상황은 모순이 되는 상황은 아니므로, 혼하지는 않다고 하더라도 얼마든지 일어날 수 있는 상황”⁸⁾이다. 따라서 홀의 반례에 대한 김성수(2012)의 해결책은 일반적인 해결책이 될 수 없고 따라서 활성화 경로 이론은 여전히 강력한 반례를 갖는 이론으로 남는다.

활성 경로 이론에 대한 또 다른 반례인 (치명적 해독제)는 어떠한가? 활성화 경로 이론에 따르면 보디가드가 치명적 해독제를 넣은 것과 암살자가 독을 넣은 것 둘 다 보스가 살아남은 것의 원인이 되는데, 만약 이러한 관정에 문제가 있다면 이는 활성화 경로 이론에 대한 반례가 된다. 김성수(2012)는 암살자가 독을 넣은 것은 보스가 살아남은 것의 원인일 수 없다는 직관과 관련하여, 보디가드가 치명적 해독제를 넣은 것만이 보스가 살아남은 것의 원인이라는 주장은 생각만큼 강한 직관적 설득력을 갖지 않는다고 주장한다.⁹⁾ 그 이유는, 해독제가 그것만 마셨을 때 치명적이라는 점을 고려해볼 때 치명적 해독제가 든 커피를 마신 것이 원인이 되어 보스가 살아남았다는 주장이 성립하기 위해서는 이 원인과 함께 작동하는 또 다른 원인(join cause)이 있어야 할 것 같기 때문이다. 그렇다면 그 해독제와 함께 작동하는 독을 암살자가 넣었다는 것을 또 다른 원인으로 보는 것이 생각만큼 반직관적이지 않다. 이 주장과 관련하여 김성수(2012)는, 치명적 해독제와 유사한 인과 구조식을 갖지만 커피에 독약을 넣는지의 여부 대신 커피에 이미 독이 들어있는지의 여부를 변수로 택한 인과 모델을 고려한다. 그리고 이 모델에 대해 활성화 경로 이론은 커피에 이미 독이 들어 있다는 것과 치명적 해독제를 넣은 것 모두 원인이라는 관정을 내리며 이는 직관과 일치

8) 김세화 (2015), p. 149.

9) 김성수 (2012), p. 284.

하는 답이라는 점을 지적한다.¹⁰⁾

김세화 교수는 이런 식의 해법은 “커피에 독이 있는 것이 보스가 살아남은 것의 원인이라는 것을 일단 받아들이면, 커피에 독이 들어 있는 것의 원인은 누군가가 커피에 독을 넣은 것이므로, 당연히 누군가가 커피에 독을 넣은 것 역시 보스가 살아남은 것의 원인”¹¹⁾이라는 주장을 하는 것으로 해석한다. 그리고 이 주장은 “인과의 전이성을 전제할 때 가능한 논변”¹²⁾인데 반해 활성화 경로 이론은 인과 전이성을 부정하므로 잘못된 것이라고 비판한다. 김세화 교수가 주장하듯 (치명적 해독제)를 비판하기 위해 인과 전이성을 전제할 수는 없다. 그러나 이 반례에 대응하는 김성수(2012)의 전략은 인과 전이성을 전제하는 것이 아니라, 먼저 강력한 반례인 (홀의 사례)를 비판하여 활성화 경로 이론의 타당성을 확보하고 그것에 기대어 (치명적 해독제)에 대한 직관을 조정하고자 하는 것이었다. 하지만 (홀의 사례)에 대한 기존의 해결책이 성립하지 않으므로 이 전략이 작동하지 않는다는 점은 분명하다.

앞서 언급했듯 (치명적 해독제)에 대한 기존의 해결책에서 핵심이 되는 아이디어는 보디가드가 치명적 해독제를 넣은 것과 암살자가 독을 넣은 것이 함께(jointly) 작동하여 그 결과로 보스가 살아남는 것으로 볼 수 있다는 것이다. 반면 김세화 교수에 따르면 “(치명적 해독제)에서는 보디가드가 커피에 해독제를 넣는 것만이 원인”이며 “설사 (치명적 해독제)에서도 보디가드가 커피에 넣은 해독제는 그것만 먹으면 치명적이므로 이것 이외에 또 다른 원인이 필요하다는 것을 인정하더라도, 여기에서 필요한 또 다른 원인은 ... ‘커피에 독이 들어있는 것’이지 ‘암살자가 보스의 커피에 독을 넣는 것’은 아닌 것으로 보인다.”¹³⁾ 이는 암살자가 커피에 독을 넣

10) 김성수 (2012), p. 285.

11) 김세화 (2015), p. 146.

12) 김세화 (2015), p. 146.

은 것은, ‘원인’을 어떤 식으로 이해하든 상관없이, 결코 보스가 살아남은 것의 원인이 될 수 없다는 주장으로 보인다.

하지만 이는 좀 더 논의할 필요가 있는 주장이다. 암살자가 커피에 치명적인 독을 넣은 상황에서, 보디가드가 커피에 치명적인 해독제를 넣지 않았더라면 보스는 죽었을 것이다. 그리고 보디가드가 커피에 치명적인 해독제를 넣은 상황에서, 암살자가 커피에 독을 넣지 않았더라면 보스는 죽었을 것이다. 보디가드가 커피에 치명적인 해독제를 넣은 것뿐만 아니라 암살자가 커피에 독을 넣은 것 역시 보스가 살아남은 것의 원인이라는 주장과 관련하여 다음의 경우를 생각해보자. 철수는 (중성의 상태인) 물에 산성 물질을 넣었고 (이와 독립적으로) 영희는 알칼리성 물질을 넣었다. 그리고 그 결과 리트머스 시험지의 색깔은 흰색을 보여주었다. 이 경우 철수가 물에 산성 물질을 넣었다는 것과 영희가 알칼리성 물질을 넣었다는 것은 각각 리트머스 시험지가 흰색을 보여주는 데 기여하는 원인이다. 물론 (치명적 해독제)는 암살자가 독을 넣는 것이 원인이 되어 보디가드가 해독제를 넣게 된다는 점에서 철수-영희의 경우와는 차이를 보인다. (치명적 해독제)에서 암살자의 행위는, 한편으로는 커피 속에 치명적 독이 들어있게끔 하면서도 다른 한편으로는 (보디가드가 커피에 해독제를 넣어) 커피 속의 독과 중화 작용을 하는 치명적 해독제가 들어있게끔 한다는 점에서, *전반적으로 볼 때* 보스의 생존 여부에 어떤 영향도 주지 못하며, 바로 이런 의미에서 보스가 살아남은 것의 원인이 될 수 없다. 필자가 보기에는, 암살자가 커피에 독을 넣은 것은 어떤 의미에서는 보스가 살아남은 것의 원인이면서 또 다른 의미에서는 보스가 살아남은 것의 원인이 아니라는 두 주장 모두를 하는 것이 가능하다.

그러나 이런 가능성을 허용한다고 해도 치명적 해독제에 대한

13) 김세화 (2015), pp. 147-148.

기존의 대응에는 여전히 문제가 있다. 그 이유는 활성 경로 이론 자체가 위와 같은 두 가지 가능성 모두를 인정하는 해석을 허용하지 않기 때문이다. 활성 경로 이론에 따를 때 (치명적 해독제)에서 보디가드가 치명적 해독제를 넣은 것과 암살자가 독을 넣은 것 둘다 보스가 살아남은 것의 원인으로 판명되는 이유는 B-D 경로와 A-D 경로가 활성 경로이기 때문이다. 그런데 활성 경로 이론은 A-D 경로가 활성 경로라는 사실을 암살자가 커피에 독을 넣은 것을 어떤 다른 원인과 함께 작동하는 원인(joint cause)으로 해석하는 대신 (치명적 해독제)에서 인과 전이성이 성립하는 것으로 해석한다.¹⁴⁾ 그리고 이는 잘못된 결론이다. 이제 활성 경로 이론을 그것의 반례로부터 옹호하고자 한다면 기존의 전략과는 다른 접근이 필요하다는 점이 분명해진다. 필자는 다음 절에서 기존의 활성 경로 이론의 범위를 제한하고 그것을 다시 확장함으로써 활성 경로 이론을 수정하는 가능성을 모색해보고자 한다.

4. 활성 경로 이론의 제한 및 확장

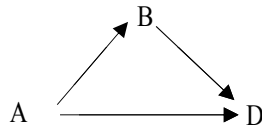
다음의 사례에 대해 생각해보자.

-
- 14) A와 B를 연결하는 경로에 속하는 유일한 경로인 A-B 경로는 활성 경로이다. 따라서 암살자가 커피에 독을 넣은 것은 보디가드가 커피에 해독제를 넣은 것의 원인이다. 그리고 B와 D를 연결하는 경로에 속하는 유일한 경로인 B-D 경로는 활성 경로이다. 따라서 보디가드가 커피에 해독제를 넣은 것은 보스가 살아남은 것의 원인이다. 만약 인과 전이성이 성립한다면, 앞의 두 가지 인과 사실로부터 암살자가 커피에 독을 넣은 것은 보스가 살아남은 것의 원인임이 도출된다. 하지만 활성 경로 이론은 인과 전이성을 전제하지 않는다. 대신 암살자가 커피에 독을 넣은 것이 보스가 살아남은 것의 원인인지의 여부는 A와 D를 연결하는 경로 중 적어도 한 경로가 활성 경로인지 여부에 달려있다. 이 사례에서 A와 D를 연결하는 경로에 속하는 A-D 경로가 활성 경로이므로, 암살자가 커피에 독을 넣은 것은 보스가 살아남은 것의 원인으로 판명된다.

(전이성) 암살자가 보스의 커피에 독을 넣었다. 이를 본 보디가드는 커피에 해독제를 넣었고 그 해독제가 커피 속의 독을 중화시켰다. (해독제는 그것만 마셨을 경우 무해하다.) 보스는 이 커피를 마셨고 살아남았다. 만약 암살자가 커피에 독을 넣지 않았더라면 보디가드는 해독제를 넣지 않았을 것이다. 만약 보디가드가 해독제를 넣지 않았더라면 보스는 커피를 마시고 죽었을 것이다.¹⁵⁾

이 사례는 인과 전이성이 성립하지 않는 경우로, 활성 경로 이론은 이 사례에 대해 정확한 답을 내놓는다.

$$\begin{aligned} A &= 1 \\ B &= A \\ D &= A \ \& \ \sim B \end{aligned}$$



여기서 변수 A는 암살자가 독을 넣는지 여부, B는 보디가드가 해독제를 넣는지 여부 그리고 D는 보스가 죽는지 여부라고 하자. B를 실제 값인 B = 1에 고정된 상태에서, A의 값을 실제 값인 A = 1에서 A = 0으로 변화시켜도 D의 값은 D = 0을 계속 유지한다. 따라서 A-D 경로는 비활성 경로이다. 마찬가지로 A-B-D 경로 역시 비활성 경로임을 쉽게 확인할 수 있다. A에서부터 D로의 활성 경로가 존재하지 않으므로, 암살자가 커피에 독약을 넣은 것(A = 1)은 보스가 살아남은 것(D = 0)의 원인이 아니라는 직관과 부합되는 결과가 도출된다.

하지만 앞서 보았듯 활성 경로 이론은 (치명적 해독제)에 대해 잘못된 답을 내놓는다. 이 점은 활성 경로 이론이 *인과 사슬에 놓인 사건들 간의 관계에 대해* 일반적으로 정확한 답을 내놓지 못한다는 점을 보여준다. 그렇다면 인과 사슬에 놓인 사건들 간의 관계

¹⁵⁾ Hitchcock (2007), p. 517. 김성수 (2012), p. 264에서 재인용.

에 대한 답을 내놓는 것 자체를 제한하는 것이 이런 문제를 해결하는 한 가지 방법이 될 것이다.

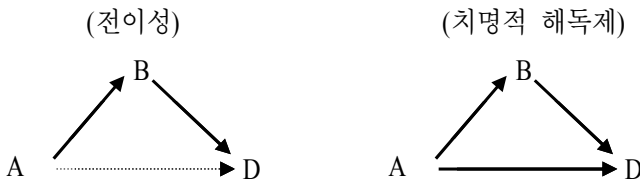
활성 경로: 주어진 경로 $\langle X, Y_1, \dots, Y_n, Z \rangle$ 는, 이 경로에 속하지 않는 모든 변수의 값을 실제 값에 고정할 때 Z 가 X 에 반사실적으로 의존하는 경우 그리고 단지 그 경우에만 활성 경로이다.

인접한 원인: (그 사이에 어떤 변수도 없이) 서로 인접해 있는 변수 쌍 $\langle C, E \rangle$ 에 대해, $C = c$ 는 C - E 경로가 활성 경로인 경우 그리고 단지 그 경우에만 $E = e$ 의 인접한 원인이다.

활성 경로 이론의 제한: 기존의 활성 경로 이론의 범위는 단지 인접한 원인을 가려내는 데로 국한된다.

제한된 형태의 활성 경로 이론은 기존의 이론과 동일하지만, 그 적용 범위를 주어진 인과 모델에서 다른 변수를 통하지 않고 직접적으로 인접해 있는 변수들 간에 성립하는 인과 관계의 파악으로 한정짓는다.

아래의 그래프는 (전이성)과 (치명적 해독제)에 등장하는 변수들 중에서 서로 인접해 있는 모든 변수들 쌍에 대해, 인접 인과 관계에 있는 쌍은 굵은 화살표를 사용하여 보여주고 그렇지 않은 쌍은 점선 화살표를 사용하여 보여준다.



기존의 이론은, (전이성)에서 A-D 간에 활성화 경로가 없다는 점을 인과 전이성이 없다는 것으로 이해하고 (치명적 해독제)에서 A-D 간에 활성화 경로가 있다는 점을 인과 전이성이 성립하는 것으로 이해했다. 하지만 이미 지적했듯 이는 활성화 경로 이론에 문제가 있음을 보여줄 따름이다. 반면 제한된 형태의 활성화 경로 이론은, (전이성)에서 A-B 경로가 인접 인과 관계에 있고 B-D의 경로 역시 인접 인과 관계에 있지만 A-D 경로는 그렇지 않다는 것을 보여준다. 그리고 (치명적 해독제)에서는 A-B, B-D, A-D 경로 모두 인접 인과 관계에 있음을 보여준다. 하지만 제한된 형태의 활성화 경로 이론은 A-B가 인접 인과 관계에 있고 B와 D가 인접 인과 관계에 있다는 바로 그 사실로 인해 A-D 역시 인과 관계에 있는지 여부에 대해서는 어떤 답도 내리지 않는다.

이런 식으로 제한된 형태의 활성화 경로 이론은 물론 그 적용 범위가 극도로 제한된 이론이다. 따라서 인접 인과 관계들로 이루어진 인과 사슬 혹은 인과 망(network)에서 성립하는 (인접 인과 관계 이외의) 인과 관계를 판명할 방법을 제시할 것이 요구된다. 이와 관련하여 생각할 수 있는 한 가지 가능성은, 인접 인과 관계로 포착되지 않는 사건들 간에 성립하는 (따라서 다른 변수의 매개를 요구하는) 인과 관계는 (다른 변수의 매개 없이) 직접적으로 성립하는 인접 인과 관계에 의해 결정되는 관계일 것이라는 점이다.

가정: 인접 인과 관계에 있지 않은 사건들 간에 성립하는 인과 관계는 인접 인과 관계에 있는 사건들로 구성된 인과 사슬 혹은 인과 망에서 이들 사건들이 갖는 어떤 관계에 의해 결정된다.

이 점을 염두에 두고, 먼저 경로들로 구성된 ‘경로-망(network)’과

경로-망에서 인접 인과 관계에 있는 변수들로만 구성된 ‘인과-망’을 정의하도록 하자.

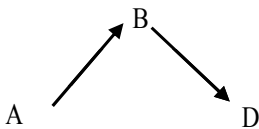
경로-망: X에서 Z로의 경로-망은 X에서부터 Z로 이어진 모든 경로로 구성된다.

인과-망: X에서 Z로의 인과-망은 X에서 Z로의 경로-망에서 인접 인과 관계에 있는 경로만을 남겨놓은 망으로 구성된다.

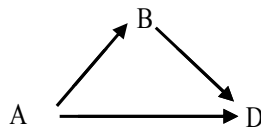
바로 이전에 논의한 (전이성)과 (치명적 해독제) 각각의 그래프에서 A에서 D로 연결되는 경로는 A-B-D와 A-D 두 경로이며 따라서 A에서 D로의 경로-망은 이 두 경로로 구성된다. 한편 각각의 그래프에서 A에서 D로의 인과-망은 점선의 화살표를 제외하고 단지 진한 실선의 화살표로만 구성된다.

이제 이런 식으로 인과-망을 구성했을 때 이들 간에 성립하는 인과 관계를 판별하기 위해 무엇이 요구되는지에 대해 생각해보도록 하자. 다음은 앞서의 두 사례에 대응하는 인과-망을 보여준다.

(전이성 인과-망)



(치명적 해독제 인과-망)



(전이성)에 해당하는 인과-망에 따르면, D와 A는 인접 인과 관계의 측면에서 서로 독립적이다. 이것이 뜻하는 바는, 적어도 인접 인과 관계에 있는 변수들 간의 의존 관계를 살펴보는 한, D는 A에 더 이상 의존하지 않는다는 것이다. 이제 이 사실과 앞서의 가정 하에

서 (전이성)과 (치명적 해독제) 모두에 대해 정확한 답을 내놓기 위해서는 어떤 방식의 컨트롤이 요구되는지 생각해볼 필요가 있다. 그리고 이 단계에서 우리가 원하는 것은 (기존의 활성화 경로 이론과는 다른 식으로) 원인과 결과 간의 반사실적 의존 관계를 방해하는 요인들을 컨트롤하는 방식이다.

이와 관련하여 D와 A가 인접 인과 관계의 측면에서 서로 독립적이라는 사실을 D에 대한 인과 구조식을 재조정(recalibration)함으로써 반영하는 다음과 같은 방식을 고려해보자. 즉 D와 A가 인접 인과의 측면에서 서로 독립적인 경우 A는 D의 (재조정된) 인과 구조식에 더 이상 등장하지 않는 변수로 상정해보자. 그리고 남겨진 변수들 간의 의존 관계만을 고려하여 D의 인과 구조식을 다시 조정해보자. (전이성)에서 일단 보스의 생사(D)가 암살자의 행동(A)와 무관하다고 상정하게 되면 이는 보스의 생사와 보디가드의 행동(B)의 관계에 대해서도 함축하는 바가 있다. (전이성)의 원래 인과 구조식에서 D가 B에 의존하는 것은 (전이성)의 상황에서는 D가 A에도 의존하기 때문이었다. 반면 일단 D와 A가 독립적인 것으로 상정하게 되면, ((전이성)에서 보디가드가 넣는 해독제는 무해하므로) D는 B에도 의존하지 않게 될 것이며 따라서 재조정된 D의 구조식은 어떤 변수에도 의존하지 않게 된다.¹⁶⁾

16) 여기서 제안된 방식으로 보스의 생사(D)에 관한 인과 구조식을 재조정하기 위해서는, 대략적으로 말해, (인접 인과 관계에 있지 않은) 암살자의 행위가 보스의 생사에 어떤 영향도 미치지 않으면서 보디가드가 실제와 달리 행동하는 가능 세계에서 보스의 생사에 어떤 일이 생길 것인지 생각해보아야 한다. 보스의 생사가 암살자의 행위에 의해 어떤 영향도 받지 않는 세계에서 보디가드가 해가 없는 해독제를 넣지 않는다고 하더라도 보스는 여전히 살아남을 것이다. 심사자들이 지적하듯 여기서 제안하고 있는 기술적인 (technical) 해결책은 그것을 지지하는 직관적인 근거에 의해 정당화될 필요가 있다. 이와 더불어 재조정과 관련된 정확한 의미론 역시 제공될 필요가 있다. 이들 문제들에 대한 논의는 이 논문의 범위를 벗어나며 앞으로 보강되어야 할 과제이다.

이렇듯 다른 매개 변수 없이 직접적으로 관련되어 있는 인접 인과 관계만으로 구성된 인과-망을 특징짓기 위해 기존의 구조식을 조정해보자.

재조정: 기존의 인과 구조식에서 A_1, \dots, A_n 에 의존하는 변수 B 에 대해, B 의 인접한 원인이 아닌 모든 A_i 들을 제거하고 남은 변수들과 B 간에 성립하는 의존 관계를 반영하여 B 의 인과 구조식을 재조정하라.

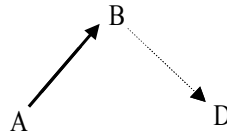
이러한 재조정 과정을 거친 인과-망을 ‘재조정된 인과-망’이라고 하자.

재조정된 인과-망: 재조정된 인과-망은 재조정된 인과 구조식을 보여주는 인과-망이다.

다음은 (전이성)에 대한 재조정된 인과-망을 보여준다.

(전이성의 재조정된 인과-망)

$$\begin{aligned} A &= 1 \\ B &= A \\ D^* &= 0 \end{aligned}$$



D 와 A 를 독립적인 것으로 상정할 때, D 는 B 에도 의존하지 않고 (D 와 B 사이에 있는 점선의 화살표가 이 점을 보여준다) 따라서 D 는 어떤 변수에도 의존하지 않는다. D^* 는 다른 어떤 변수에도 의존하지 않으므로 그것이 실제로 갖는 값을 부여받는다. 재조정된 구조식 D^* 는 이 점을 보여준다. 정리해서 말하면, 재조정되기 이전의

인과-망과 재조정된 인과-망은 다음을 제시한다. 재조정 이전의 인과-망에서 A가 D에 영향을 미칠 수 있는 유일한 방법은 B를 경유하는 것이다. 하지만 일단 재조정을 하고 나면 B 역시 D로부터 차단되므로 A는 D에 어떤 인과적 영향도 미칠 수 없다. 즉 암살자가 커피에 독을 넣은 것은 보스가 살아남은 것의 인접한 원인도 아니고 B를 경유하는 원인도 아니다. 그리고 이는 정확한 답이다.

이상의 논의는 인접 인과 관계에 있지 않은 변수들 간의 인과 관계를 판별할 방법을 제시한다.

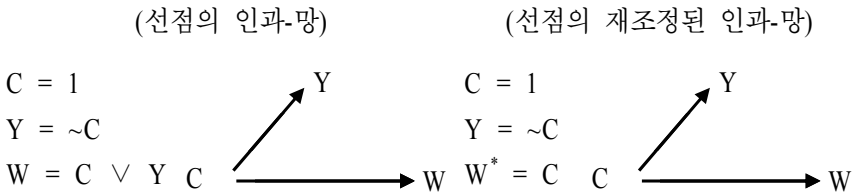
경유한 원인: X에서 Z로의 재조정된 인과-망에서 인접 인과 관계에 있지 않은 X와 Z에 대해, $X = x$ 는 (이 인과-망에 속해있지 않은 변수를 그것의 실제 값에 고정한 상태에서) Z가 X에 반사실적으로 의존하는 경우 그리고 단지 그 경우에만 $Z = z$ 의 경유한 원인이다.

경유한 원인은 특정 경로 대신 (재조정된) 인과-망 전체를 살펴볼 것을 요구한다는 점에서 기존의 활성화 경로 이론과 차이점을 보인다. (전이성)의 재조정된 인과-망에서 A가 어떤 다른 변수를 경유하여 D에 인과적 영향을 미칠 수 있는지 여부는 A의 값을 실제 값인 $A = 1$ 에서 $A = 0$ 로 변화시킬 때 D의 값 역시 변화하는지에 달려있다. 그런데 재조정된 인과-망에서 D는 다른 어떤 변수에도 의존하지 않으므로 $A = 1$ 은 $D = 0$ 의 경유한 원인이 아니다.

반면 (치명적 해독제)의 인과-망에서, 암살자가 독을 넣은 것은 보디가드가 치명적 해독제를 넣은 것과 보스가 살아남은 것의 인접한 원인이며 또한 보디가드가 치명적 해독제를 넣은 것은 보스가 살아남은 것의 인접한 원인이다. 그렇다면 암살자가 독을 넣은 것은 보스가 살아남은 것의 인접한 원인이라는 사실에 더해 경유한

원인이라는 사실 역시 성립하는가? 이에 답하기 위해서는 인과-망과 재조정에 대해 살펴봐야한다. 그런데 (치명적 해독제)의 인과-망에서 D는 여전히 동일한 변수들에 의존하고 있으므로 기존의 의존관계에 어떤 변화도 없고 따라서 재조정이 필요하지 않다. 따라서 (치명적 해독제)의 (아무 변화가 없는) 재조정된 인과-망에서는 $A = 1$, $B = A$ 그리고 $D = (A \ \& \ \sim B) \vee (\sim A \ \& \ B)$ 가 여전히 성립한다. 이제 A의 값을 $A = 0$ 로 변화시키면 $B = 0$ 이 되고 D는 아무 변화 없이 $D = 0$ 에 머문다. 즉 D가 A에 반사실적으로 의존하지 않으므로, (치명적 해독제)에서 암살자가 커피에 독을 넣은 것 ($A = 1$)은 보스가 살아남은 것($D = 0$)의 경유한 원인이 아니다. 그리고 이는 정확한 답이다.

이제 남은 것은 이런 해결책이 (선점)과 (홀의 사례)에도 적용되는지를 살펴보는 일이다. (선점)은 철수가 돌을 던져 유리창이 깨졌지만 만약 철수가 돌을 던지지 않았더라면 영희가 돌을 던져 유리창이 깨지게 되는 경우이다.



먼저 (선점)의 인과-망에 대해 살펴보자. 왼쪽 그래프 좌측의 구조식은 원래의 구조식을 보여준다. (이 글의 맨 처음에 나온 원래의 그래프에는 Y에서 W로 이어지는 화살표가 있었음을 기억하라.) (선점)에서 C-Y 경로가 활성화 경로임은 분명하며 따라서 $C = 1$ (철수가 돌을 던진 것)은 $Y = 0$ (영희가 돌을 던지지 않은 것)의 인접한 원인이다. 또한 Y를 실제 값인 $Y = 0$ 에 고정할 때, C-W 경로

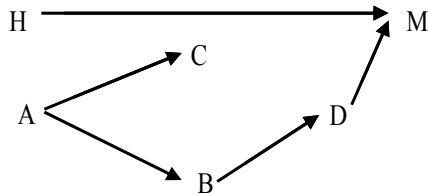
역시 활성화 경로이며 따라서 $C = 1$ 은 $W = 1$ (유리창이 깨진 것)의 인접한 원인이다. 한편 C 를 실제 값인 $C = 1$ 에 고정한 상태에서 Y 를 실제 값인 $Y = 0$ 에서 $Y = 1$ 로 변화시켜도 W 의 값은 아무 변화 없이 $W = 1$ 에 머문다. 따라서 $Y = 0$ 은 $W = 1$ 의 인접한 원인이 아니며 위의 인과-망 그래프는 Y 와 W 사이의 화살표를 제거함으로써 이를 보여준다.

(선점)의 인과-망에서 W 는 Y 와 독립적이다. 이제 남아 있는 C 와 W 간에 어떤 관계가 성립하는지 재조정을 해보자. 유리창이 깨지는지의 여부가 영희가 돌을 던지는지 여부와 무관한 것으로 상정된 경우에도, 유리창이 깨지는지의 여부는 여전히 철수가 돌을 던지는지 여부에 의존한다. 따라서 재조정된 구조식($W^* = C$)에서 W 는 여전히 C 에 의존한다. 그리고 C 와 W 사이에는 어떤 다른 변수도 존재하지 않으므로 C 와 W 는 인접 인과 관계를 갖지만 경유 인과 관계는 갖지 않는다. 이는 정확한 답이다.

마지막으로 (홀의 사례)를 살펴보자. 다음은 이 사례의 기존 인과 구조식과 인과-망을 보여준다.

(홀의 사례의 인과-망)

- $H = 1$
- $A = 1$
- $B = A$
- $C = A$
- $D = C \ \& \ \sim B$
- $M = H \ \& \ \sim D$



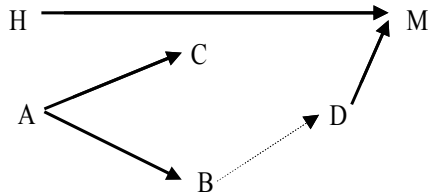
앞서의 논의에서 밝혔듯, C 를 실제 값인 $C = 1$ 에 고정한 상태에서, B 의 값을 1에서 0으로 변화시키면 D 의 값은 0에서 1로 변화

되므로 $B = 1$ 은 $D = 0$ 의 인접한 원인이다. 한편 B 를 실제 값인 $B = 1$ 에 고정된 상태에서, C 의 값을 1에서 0으로 변화시키면 D 의 값은 아무 변화 없이 0에 머문다. 따라서 $C = 1$ 은 $D = 0$ 의 인접한 원인이 아니며 따라서 C 와 D 사이를 연결하는 화살표는 없다.

기존 경로-망에 비교할 때 이 인과-망의 유일한 변화는 D 가 C 와 독립적이 되었다는 점이다. 따라서 D 와 B 간의 의존 관계를 재조정하기만 하면 재조정된 인과-망을 얻게 된다. 홀의 사례에서 d -뉴런은 c -뉴런에 의해 자극되고 b -뉴런에 의해 억제된다. 이제 D 가 C 와 독립적이라는 사실은 이 인과-망에서 D 의 값이 $D = 1$ 이 될 가능성을 제거한다. 그렇다면 B 의 값이 무엇이든 상관없이 D 는 1의 값을 가질 수 없고 D 는 B 와 독립적이 된다. 따라서 이 인과-망에서 D 는 어떤 변수에도 의존하지 않는 변수이며 D 에게는 그것이 실제로 갖는 값인 $D^* = 0$ 이 부여되어야 한다.

(홀의 사례의 재조정된 인과-망)

- $H = 1$
- $A = 1$
- $B = A$
- $C = A$
- $D^* = 0$
- $M = H \ \& \ \sim D$



재조정된 인과-망에서 A 가 D 에 어떤 영향도 미칠 수 없음이 자명하다. A 의 값을 어떤 식으로 변화시켜도 D 의 값은 0에 머물기 때문에 $A = 1$ 은 $D = 0$ 의 (인접한 원인이 아닐 뿐만 아니라) 경유한 원인도 아니다. 그리고 이는 정확한 답이다. 이와 마찬가지로, M 의 값은 H 와 D 의 값에 의해 결정되므로 A 는 M 에 영향을 미칠 수 없

고 따라서 $A = 1$ 은 $M = 1$ 의 경유한 원인이 아니다. 그리고 이는 정확한 답이다. 이상의 방법으로 활성 경로 이론을 제한하고 확장하면 기존의 반례들을 다룰 수 있는 가능성이 확보된다.

5. 맺음말

활성 경로 이론은 주어진 상황에서 결과와 원인 간의 반사실적 의존 관계를 방해하는 요인들을 컨트롤하는 방식을 제공함으로써 인과 관계를 반사실적 의존 관계로 분석하고자 하는 이론이다. 하지만 이 이론 역시 반례를 갖는다. 김세화 교수는 이러한 반례로부터 활성 경로 이론을 옹호하고자 하는 기존의 시도를 설득력 있게 비판하였다. 이 비판에 비추어 볼 때 활성 경로 이론은 여전히 심각한 반례를 갖는 이론이라는 점을 받아들이거나 혹은 그 반례들을 극복할 수 있는 또 다른 가능성을 모색해야 한다. 그 가능성 중의 하나는 두 단계에 걸친 제한과 확장이다. 그 첫 번째 단계는 활성 경로 이론의 범위를 인접 인과 관계의 규명으로 제한하는 것이다. 그리고 두 번째 단계는 제한된 형태의 활성 경로 이론이 밝혀낸 인접 인과 관계들로부터 경유 인과 관계를 규명할 수 있는 방법을 제공하는 것이다. 이 방법은 인과 관계 파악을 위해 각 단계별로 서로 다른 컨트롤 방식을 요구하므로 기존의 이론보다 더 복잡하다. 하지만 두 단계 모두 인과 관계를 (적절한 컨트롤 하에서의) 반사실적 의존 관계로 파악하므로 반사실적 인과론의 핵심을 일관성 있게 유지한다. 그리고 이런 식으로 제한되고 확장된 형태의 활성 경로 이론은 기존 이론에 대한 반례를 극복할 수 있다. 마지막으로 이러한 시도가, 김세화 교수가 강조하듯, 활성 경로 이론 혹은 인과 이론 전반에 대한 재점검과 비판적 논의에 기여할 수 있길 바란다.

참고문헌

- 김성수 (2012), “반사실적 인과론과 인과 구조식 접근법: 활성화 경로 이론의 재검토”, 『철학』 113, pp. 259-291.
- 김세화 (2015), “김성수 교수의 활성화 경로 이론에 대한 변호와 그에 대한 반론”, 『논리연구』 18 (1), pp. 133-152.
- Hall, Ned (2007), “Structural Equations and Causation”, *Philosophical Studies* 132, pp. 109-136.
- Halpern, J. and Pearl, J. (2005), “Causes and Explanations: A Structural-Model Approach Part 1: Causes”, *British Journal for the Philosophy of Science* 56, pp. 843-887.
- Hitchcock, Christopher (2001), “The Intransitivity of Causation Revealed in Equations and Graphs”, *Journal of Philosophy* 98, pp. 273-299.
- Lewis, David (1973), “Causation”, *Journal of Philosophy* 70, pp. 556-567.
- Lewis, David (1986), “Postscripts to “Causation””, in D. Lewis, *Philosophical Papers*, Vol. 2, New York: Oxford University Press, pp. 172-213.
- Lewis, David (2000), “Causation as Influence”, *Journal of Philosophy* 97, pp. 182-197.
- Menzies, Peter (2004), “Causal Models, Token Causation, and Processes”, *Philosophy of Science* 71, pp. 820-832.

서울시립대학교 철학과

Department of Philosophy, University of Seoul

sungsukim@uos.ac.kr

The Active-Route Account Restricted and Expanded:
A Reply to Seahwa Kim's Criticisms

Sungsu Kim

The idea that an effect counterfactually depends on its cause is simple and intuitive. However, this simple idea runs into various difficulties. The active route account, in order to avoid the difficulties, analyzes causation in terms of counterfactual dependence under certain control. In her recent article, Seahwa Kim criticizes Sungsu Kim's earlier attempt to defend the active route account from its counterexamples. Her criticisms are convincing, and defenders of the active route account or counterfactual analysis of causation in general need another defense. In response, a two-step defense is proposed. First, the scope of the active route account is restricted to 'proximate' causal relation. Second, a control over factors that are in proximate causal relation is offered to figure out 'distant' causal relation. The result is that with proper control, an effect indeed counterfactually depends on its cause.

Key Words: Causation, Counterfactual theories of causation, Active route account, Counterexample, Seahwa Kim