

인과의 조종가능성 이론과 간섭*†

여 영 서

【요약문】 현장의 과학자들이 지닌 인과 개념을 가장 잘 반영하고 있는 것으로 보이는 인과의 조종가능성 이론은 인과 관계를 조종가능성으로 이해할 수 있다는 직관에 기초한다. 즉 C 가 E 의 원인이라는 것은 C 를 조종함으로써 E 를 변화시킬 수 있다는 것이다. 이 이론의 핵심 개념인 조종의 이해를 위해 최근 관심의 초점이 되는 간섭은 이상화된 실험 상태에서 국지적 변화를 일으키는 것이라고 이해할 수 있다. 우드워드는 간섭의 의미가 목적에 따라 다르게 구체화될 수 있다는 전제 아래 자신이 제시한 간섭 개념이 필의 간섭 개념과 다르다고 주장한다. 하지만 본 논문은 이와 같은 우드워드의 주장이 설득력이 없다고 결론짓는다. 본 논문은 우드워드의 간섭 개념이 필의 간섭 개념과 다르기 보다는 필의 간섭 개념이 암묵적으로 전제하고 있는 바를 구체화하고 있다고 논증한다.

【주제분야】 인과, 과학철학, 형이상학,

【주제어】 인과, 조종가능성, 간섭, 필, 우드워드

* 접수완료: 2006. 12. 20 / 심사 및 수정완료: 2007. 2. 10

† 본 논문은 2005년도 한국학술진흥재단의 지원(KRF-2005-079-AS0034)에 의해 연구되었음. 본 논문의 초고는 2006년 논리학회 6월 학회에서 발표되었음. 본 논문에 대해 날카로운 지적과 좋은 제안을 해 주신 본 학회지의 심사위원들께 감사드린다. 또 본 논문의 준비 과정에서 좋은 조언을 주신 이초식 교수님 및 “인과의 철학적 분석” 연구팀 선생님들, 그리고 김동현 씨와 박일호 씨에게 감사드린다.

1. 들어가기

인과의 조종가능성 이론은 인과 관계를 조종가능성으로 이해할 수 있다는 직관에 기초한다. 즉 C 가 E 의 원인이라는 것은 C 를 조종함으로써 E 를 변화시킬 수 있다는 것이다. 이러한 직관은 현장의 과학자들이 사용하는 인과 개념에 잘 들어맞는다. 따라서 현장을 존중하는 철학적 작업의 필요성에 동의한다면, 인과의 조종가능성 이론에 대한 철학적 관심의 필요성에도 동의할 것이다.¹⁾

실험을 통해 인과 관계를 입증하려는 현장 과학자들의 작업은 대체로 적절하게 주어진 조건 아래에서 원인이라고 생각되는 것을 작동시켜 조종함으로써 원하는 결과를 가져오려는 시도라고 할 수 있다. 이에 원인을 결과를 낳는 방편 또는 도구로 이해하는 인과의 조종가능성 이론은 통계학자들, 사회과학자들, 그리고 자연과학자들이 인과를 이해하고 사용하는 방식을 잘 드러낼 수 있다. 또 바로 그 점은 현장의 과학자들이 인과의 조종가능성 이론을 높이 평가하는 이유이기도 하다.

하지만 많은 철학자들은 조종 개념에 의해 인과를 이해하려는 시도가 순환적이라는 비판과 인간의 행위를 통해 조종할 수 없는 인과 관계를 어떻게 설명할 것인가라는 비판 등을 제기하며, 인과의 조종가능성 이론이 심각한 난관에 봉착하여 있다고 판단한다. 그렇다면 인과와 조종의 연관 관계를 탐구할 때 인과의 본성을 이해할 수 있다는 비철학자들의 접근 방식이 잘못된 것인가? 아니면

1) 멘지스와 프라이스는 조종가능성 이론이 철학자들 사이에 인기가 없는 이유로 두 가지를 제시한다. 첫째, 조종가능성 이론의 장점이 철학자들에게 충분히 알려지지 못했고, 둘째, 그 이론의 문제점으로 지적된 것들이 극복 불가능하다는 잘못된 판단이 널리 퍼져 있었다는 것이다. (Menzies & Price, p. 187)

철학자들의 비판이 적절하지 못한 것인가? 철학자들이 지적하고 있는 인과의 조종가능성 이론의 문제점은 극복하기 어려운 것인가? 또 그런 문제점은 인과의 조종가능성 이론이 인과의 본성을 이해하는 데에 아무런 도움이 되지 못한다는 결론의 충분한 근거가 되는가?

철학자들과 비철학자들 사이에 보이는 상반된 태도를 이해하기 위해 우리는 인과 이론에 관심을 가지는 철학자들과 비철학자들의 목적이 서로 다르다는 점에 주목할 필요가 있다.²⁾ 환원주의를 옹호하는 철학자들은 인과의 본성을 이해하기 위해 인과 개념이 전혀 포함되지 않은 다른 개념을 이용해서 인과를 설명할 수 있어야 한다고 본다. 이와 같은 환원주의적 접근 방식에서 볼 때 조종이라는 개념은 인과 개념을 분석하기에 부적절할 수 있다. 왜냐하면 조종한다는 것은 변화하도록 만든다는 인과의 의미를 어느 정도 포함하는 개념이기 때문이다. 하지만 환원주의적 접근 방식을 취하지 않는 철학자들과 대부분의 비철학자들에게 순환성 그 자체는 큰 문제가 아니다. 인과를 조종에 의해 환원하지 못하더라도 조종이라는 개념이 인과의 본성을 밝히는 데에 도움이 된다면, 인과와 조종이라는 두 개념이 어떻게 연결되어 있는지를 분석하는 조종가능성 이론은 그럴듯한 작업이라고 평가할 수 있다는 것이다.

본 논문에서 논자는 조종(manipulation)이라는 개념이 인과의 본성을 밝히는 데에 도움이 된다고 전제하고, 이러한 전제 아래 조종이라는 개념을 어떻게 이해할 것인가라는 점에 초점을 맞춘다. 특히 본 논문은 간섭(intervention) 개념을 새로 제시하며 조종 개념을 이해하려는 최근의 논의 중 우드워드(J. Woodward)가 제시한 간섭 개념과 펄(J. Pearl)이 제시한 간섭 개념 사이에 어떤 차이점이 존재하는지를 검토해 볼 것이다. 이러한 논의를 통해 논자

2) 이 점은 우드워드도 지적하고 있다. (Woodward, 2001, p. 2)

는 인과의 이해를 위해 인과의 조종가능성 이론이 제시하는 바가 무엇인가에 대해 생각해보고자 한다.

2. 인과의 조종가능성 이론

개스킹(D. Gasking), 콜링우드(R. Collingwood), 폰 라이트(G. von Wright), 멘지스(P. Menzies), 프라이스(H. Price), 우드워드 등에 의해 발전된 인과의 조종가능성 이론은 인과의 행위자 이론(the agency theory of causation) 또는 실험주의자의 인과 이론(the experimentalist's theory of causation)이라고 불리기도 한다.

조종가능성 이론의 고전적 입장을 제시하여 철학적으로 많은 주목을 받았던 폰 라이트는 특히 인간의 가능한 행위라는 측면에 초점을 맞춰 인과 관계를 분석한다. 그의 정의에 따르면, p 가 q 의 원인이라는 말의 의미는, p 를 산출(produce)한다면 q 를 일으킬 것(bring about)이라는 것이다. 이때 원인에 해당하는 p 를 산출하는 것은 인간의 행위이다.³⁾

3) 폰 라이트는 이와 같은 인과의 정의를 기초로 한 조종가능성 이론이 성립하기 위해 다음의 세 가지 전제가 충족되어야 한다고 주장한다. 첫째, 조종가능성 이론은 행위자가 간섭하려고 할 때 그가 어떤 상황에 처해 있는지에 대해서 잘 알고 있어야 한다고 요청한다. 그것은 만약 행위자가 간섭하지 않는다면 세계가 어떻게 돌아갈지에 대한 지식을 포함한다. 둘째, 조종가능성 이론은 논리적 원자주의를 전제한다. 이 때 말하는 논리적 원자주의는 인과적 연관이 있는 각각의 사태가 개념적으로나 검증주의적으로 분리될 수 있다는 것이다. 즉 원인이라고 지적되는 사건의 발생과 결과라고 지적되는 사건의 발생이 서로 독립적일 수 있어야 한다. 셋째, 조종가능성 이론은 행위자가 원인과 결과 사이에 나타나는 상례성을 파악하고 있다는 전제를 가진다. (von Wright, p.119)

폰 라이트와 마찬가지로 세계의 변화에 영향력을 행사할 수 있는 행위자의 행위에 초점을 맞춘 멘지스와 프라이스의 행위자 이론에 따르면, p 라는 사건이 q 라는 다른 사건의 원인이라는 것은 자유 의지를 가진 행위자가 q 를 발생시키기 위한 효율적 방편으로 p 를 산출하는 행위를 수행하는 것이다. 즉 p 가 q 를 야기한다는 말은 $Pr(q | p) > Pr(q | \neg p)$ 로 정의된다. 자유 의지를 가진 행위자가 q 를 발생시키기 위한 방편으로 p 를 산출하지 않기보다는 p 를 산출하는 것이 효율적이라는 것이다. 자유 행위와 인과의 연관을 통해 원인이 결과의 확률값을 높인다는 인과의 확률적 분석을 옹호하는 인과의 행위자 이론은 폰 라이트의 이론을 비결정론적 세계에서도 적용할 수 있도록 수정한 것으로 볼 수 있다.⁴⁾

그러나 폰 라이트가 제시하는 인과 개념뿐만 아니라 멘지스와 프라이스가 제시하는 인과 개념에서 조종가능성 이론의 문제점으로 지적된 것들을 극복하기에 충분하지 않다는 것이 대체적인 판단이다. 그와 같은 문제점을 하우스만(D. Hausman)은 다음 세 가지로 정리한다. 첫째, 불합리하게 제한적인 설명 영역의 문제, 즉 인간의 조종이 가능하지 않다면 인과 관계가 성립하지 않는다는 문제, 둘째, 인과를 정의함에 있어 '산출한다'거나 '일으킨다' 등의 개념을 이용한다는 점에서 순환적이라는 문제, 셋째, 인과 관계가 오로지 인간의 시각에 의존하는가라는 의인화(anthropomorphism) 문제 등이다.⁵⁾

이와 같은 문제점에 부딪혀 만족스러운 진전을 이루지 못하던 인과의 조종가능성 이론은 등식과 방향 그래프(directed graph)를 사용하여 인과 관계를 분석하고 구체적으로 응용하는 작업에 초점

4) 멘지스와 프라이스는 인과 개념이 색깔 개념과 마찬가지로 이차 성질이라고 주장하며, 색깔 개념에 대한 현대의 철학적 논의에 비유하여 인과의 행위자 이론을 옹호한다. (P. Menzies & H. Price, 1993.)

5) Hausman, p. 89.

을 맞춘 새로운 접근 방식에 의해 관심을 끌고 있다. 이 과정에서 도입된 간접 개념은 대략 이상화된 실험 상태에서의 조종으로 이해할 수 있다. 즉 간접은 X 가 Y 의 원인인지를 판단하는 데에 이용하기 위해 X 를 조종하는 행위이다. 이제 간접 개념의 좀 더 엄밀한 이해를 위해 등식과 방향 그래프를 사용한 인과 관계의 분석이 어떻게 이루어지고 있는지를 검토하기로 한다.

3. 등식과 방향 그래프

등식과 방향 그래프를 사용하여 인과 관계를 분석하는 작업은 인과 관계의 연구에 있어 최근 일반화된 접근 방식이다. 펄은 그 장점을 세 가지로 제시한다.⁶⁾ 그것들은 첫째, 중요한 가정을 표현하기가 용이하다는 것, 둘째, 결합 확률 함수(joint probability functions)를 경제적으로 표현할 수 있다는 것, 셋째, 관찰로부터의 추론이 효과적이라는 것이다.

등식과 방향 그래프를 통한 인과 관계 분석이 구체적으로 어떻게 사용될 수 있는지를 하나의 예를 통해 알아보자. 한국인의 1년 술 소비량이 절반으로 준다(A)고 할 때 나타나는 양주 및 원액 수입의 절감 효과(B)와 국민 건강의 증진 효과(C)가 국가 경제에 미치는 효과(D)가 무엇인가에 관한 관계는 다음의 등식으로 제시할 수 있다.

$$(1) B=aA$$

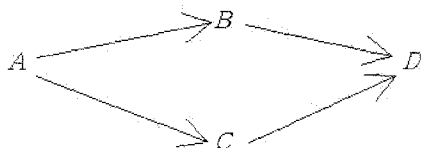
$$(2) C=bA$$

$$(3) D=cB+dC$$

6) Pearl, p. 13.

이때 (1)-(3)의 등식에서 등식 오른쪽에 있는 변수 A, B, C 등은 등식 왼쪽에 있는 변수의 직접 원인(direct cause)으로 해석된다.⁷⁾ (1)-(3)의 등식 체계는 관심의 대상인 시스템 속의 변수들 사이의 직접 인과 관계를 모두 기술하고 있다. 각각의 등식은 특정한 인과 관계를 나타낸다. 즉 각각의 등식은 자율적(autonomous)이다.⁸⁾

(1)-(3)의 등식은 다음의 방향 그래프를 통해 변수 A, B, C, D 사이의 인과 관계를 나타낼 수 있다. 이 때 화살표는 직접 인과 관계를 나타낸다. 즉 아래의 방향 그래프에서 A 로부터 B 로의 화살표는 A 가 B 의 직접 원인(direct cause)이라는 것을 나타낸다.



- 7) 본 논문에서는 전체 원인(total cause)과 기여 원인(contributing cause)을 구분하지 않고 원인이라는 용어를 사용한다. X 가 Y 의 전체 원인이라고 하는 것은 X 의 영향 전체가 Y 에 나타나는 것 즉 X 에 가해지는 단 하나의 조종만으로도 Y 의 값을 변화시킨다는 것 오직 그 경우뿐이다. 따라서 X 에 간섭하여 Y 값을 변화시키는 것은 X 가 Y 의 전체 원인이라는 것의 필요충분 조건이 된다. 기여 원인은 전체 원인과 대조되는 것으로 전체 원인이 되지 못하는 여전히 결과의 변화를 일으키는 것이다.

다음 인과 구조의 방향 그래프를 생각해보자.



여기에서 X 는 Y 의 직접 원인이고 동시에 Z 의 직접 원인이다. 그리고 Z 는 Y 의 직접 원인이다. 그런데 Y 는 X 와 Z 의 영향을 동시에 받고 있다. 따라서 X 는 Y 의 전체 원인이 아니라 기여 원인이 된다. X 가 Y 의 기여 원인이라는 것은, X 를 조종하면 Y 의 값이 변화하지만 Y 의 값이 변화했다고 해서 X 의 값의 변화를 추정할 수 없다는 것이다.

- 8) 각각의 등식이 자율적이라는 의미는 각 등식을 다른 등식에 영향을 주지 않고 간섭하여 변형시킬 수 있다는 뜻이다.



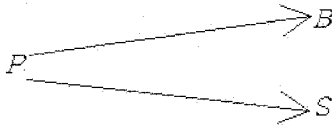
비록 방향 그래프는 인과 관계를 추론하는데 유용할 수 있지만 등식만큼 인과 관계를 완전히 나타내지 못한다. 예를 들어 $R=aP+bQ$ 와 $F=S \cdot O$ 를 비교해보자. 두 등식의 방향 그래프는 모두 같은 구조를 가진다.

방향 그래프는 $F=S \cdot O$ 에서처럼 S 와 O 가 서로 상호작용을 할 때 F 에 영향을 끼친다는 점을 적절히 나타내지 못한다. 즉 $R=aP+bQ$ 에서는 P 의 값을 고정시키고 Q 의 값을 변화시킬 때 그 결과 R 의 값이 변화된다. 하지만 $F=S \cdot O$ 에서는 S 의 값을 고정시키고 O 의 값을 변화시킬 때 F 의 값이 변화하는 경우도 있지만 변화하지 않는 경우도 있다. 등식은 이러한 차이를 나타내주지만 방향 그래프는 이 차이를 보여주지 못한다. 그렇지만 방향 그래프는 등식보다 더 쉽게 인과 관계를 파악할 수 있게 해준다. 따라서 방향 그래프와 등식은 필요에 따라 사용할 수 있으며, 두 방식을 서로 보완하며 사용하면 인과 관계를 정확하게 분석하는 데에 유용하다.

이제 간섭 개념이 만족시켜야 하는 조건이 무엇인지를 방향 그래프를 사용하여 간단히 살펴보자.⁹⁾ 일단 간섭을 I 로 표기하고, 인과 관계와 상관 관계를 구분함에 있어 I 가 만족시켜야 할 조건을 공통 원인의 문제를 통해 살펴보자.

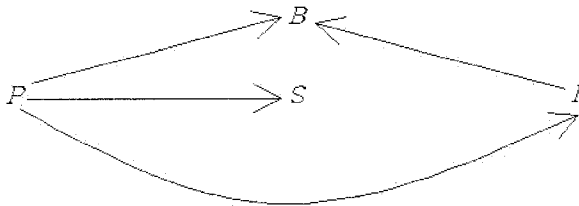
P 는 B 와 S 의 공통 원인인 앞의 예를 다시 생각해 보자.

9) Woodward (2001) pp. 95-98.

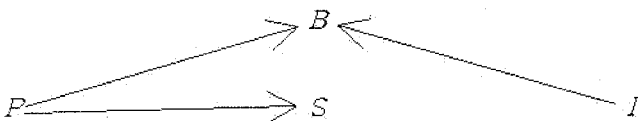


B 와 S 가 인과 관계에 놓여 있는지를 확인하기 위해, 우리는 둘 중 하나를 조종하여 다른 하나가 그에 상응하는 변화를 보이는지를 살펴봐야 한다. B 가 S 의 원인인지를 확인하고자 한다면 우리는 먼저 B 를 조종해야 할 것이다.

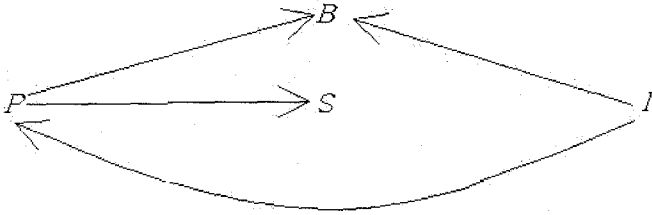
만약 B 를 조종하기 위한 간섭 I 가 P 의 변화를 반영하여 B 를 변화시키고 있다면, S 역시 B 의 변화에 상응하는 변화를 보일 것이고, 둘 사이에는 여전히 상관 관계가 유지되기 때문에 우리는 둘 사이에 인과 관계가 성립한다고 말할 것이다. 사실 S 의 변화는 B 의 변화가 원인이 아니라 P 의 변화를 따른 것이기 때문에 B 와 S 사이에는 인과 관계가 성립하지 않지만 말이다.



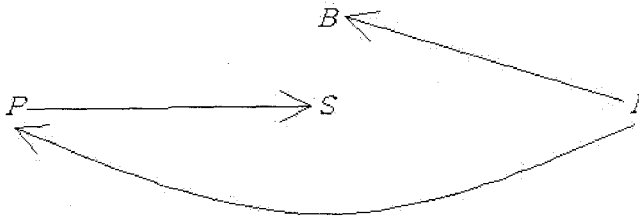
이 문제를 해결하기 위해 I 가 원인을 가지지 않는 것이라고 규정하자. 그것은 B 를 조종하는 I 가 P 의 변화를 반영하지 못하도록, 즉 I 로 들어오는 화살표가 없도록 만드는 것이다.



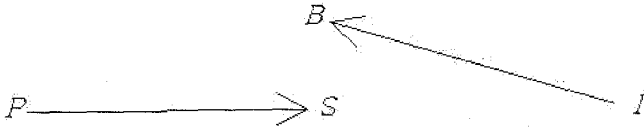
또 I 가 P 의 변화를 반영하지 않더라도 거꾸로 I 가 B 에 간섭하면서 동시에 P 의 변화를 일으킨다면 B 는 여전히 S 와 상관 관계를 유지하게 된다.



즉 B 를 조종하는 과정에서 B 를 통하지는 않지만 다른 경로를 거쳐 S 의 변화를 일으키는 것이 가능하고, 바로 그 경우 B 에 간섭하는 I 는 B 로 하여금 S 와 상관 관계를 계속 유지하게 한다. 이런 문제점을 막기 위해 간섭을 할 경우 간섭을 받는 것은 어떤 다른 원인도 가지지 못하도록 해야 한다. 즉 앞의 예에서 B 를 향한 화살표는 I 이외의 어떤 것도 허용되지 않는 것이다.



위의 그래프에서처럼 I 로부터 P 로 가는 화살표가 계속 남아있다면 I 가 B 에 간섭하여 변화를 일으키고 또 P 를 변화시켜 S 에 변화를 일으키면 여전히 B 와 S 는 상관 관계를 유지하게 되고, 둘 사이에는 인과 관계가 성립한다는 판단을 하도록 한다. 따라서 I 는 간섭을 받는 것이 어떤 다른 원인도 가지지 못하도록 해야 할 뿐만 아니라 간섭에 의해 검사하려는 인과 관계의 결론 부분에 있는 것을 다른 경로를 통해 일으키면 안된다.



인과 관계를 확인하는 실험이 성공하기 위해서는 적어도 이상에서 제시된 조건들을 만족시키는 간섭이 이루어져야 한다. 이상의 논의를 염두에 두고 이제 간섭 개념을 더 엄밀하게 정의하려는 시도를 검토하기로 한다.

4. 펄과 우드워드의 간섭

간섭을 정확하게 정의하려는 시도가 여럿 있지만,¹⁰⁾ 대체로 간섭은 이상화된 실험 상태에서 국지적(surgical) 변화를 일으키는 것으로 규정된다. 이 때 국지적 변화는 결과에 어떤 변화가 나타났다면 그 변화는 원인이라고 지적되는 것과의 인과 관계에 따른 것으로 해석할 수 있도록 확인시켜 준다.¹¹⁾

펄은 인과적 의존이라는 질적 개념을 원초 개념으로 삼아 간섭의 개념을 정의하고, 다시 간섭 개념을 기초로 하여 다양한 양적 인과 개념을 제시하고자 한다. 이 과정에서 펄은 자신이 제시한 간섭 개념을 ‘원자적 간섭(atomic intervention)’¹²⁾이라고 부른다. 펄은 국지적 변화를 일으키는 간섭의 더 엄밀한 규정을 위해 $X_i = F(Pa_i, U_i)$ 의 등식 체계로 인과 모델을 세운다. Pa_i 는 X_i 의 직

10) 다음의 작업이 대표적이다. Cartwright & Jones (1991), Spirtes et al. (1993), (2000), Pearl (1995), (2000), Woodward (1997), Hausman (1998).

11) Woodward (2001), p. 8.

12) Pearl (2000), p. 70.

접 원인이고, U_i 는 모든 배제된 변수의 영향을 요약한 오류 변수(an error variable)이다. 각각의 등식은 구분되는 인과 관계를 나타내며, 각각의 등식은 다른 인과 관계에 영향을 끼치거나 영향을 받지 않는다는 의미에서 자율적(autonomous)이다.

이제 $X_i = F(Pa_i, U_i)$ 의 인과 모델에서 변수 X_i 가 특정값 x_i 를 갖도록 하는 것이 간섭인데, 필은 X 의 값을 간섭에 의해 특정값 x_i 로 결정하는 명제를 'do' 조작사를 사용하여 $do(X_i = x_i)$ 또는 $do x_i$ 로 나타낸다. $do(X_i = x_i)$ 는 이전의 등식 $X_i = F(Pa_i, U_i)$ 대신 X_i 의 값이 x_i 가 되도록 만드는 새로운 등식으로 대체하는 것이다. 즉 X_i 의 값이 전적으로 간섭에 의해 결정되도록 기존의 등식에서 X_i 와 X_i 의 부모변수(parents)사이의 관계를 완전히 지워버리는 것이다. 따라서 $do x_i$ 는 X_i 가 좌변에 나타나는 경우, 즉 X_i 가 결과로 나타나는 모든 등식을 다음 등식 $X_i = x_i$ 로 대체하고, X_i 가 우변에 나타나는 경우, 즉 X_i 가 원인으로 나타나는 경우 X_i 의 값을 x_i 로 결정하며, X_i 가 아예 나타나지 않는 경우 등식을 그대로 두는 것이다.¹³⁾

우드워드드는 필의 간섭 개념에 대해 두 가지 문제점을 지적한다.¹⁴⁾ 첫째, 필의 간섭 개념에 따르면 X_i 가 결과로 나타나는 등식을 빼고 나머지 등식은 그대로 두어야 한다. 이 때 X_i 를 원인으로 하고 Y 를 결과로 가지는 등식이 나머지 등식 중에 포함되어 있을 수 있다. 이 경우 순환성의 문제가 나타난다는 것이 우드워드드의 지적이다. X_i 가 Y 의 원인인지를 따지기 위해 X_i 에 간섭을 하는 것인데 X_i 를 원인으로 하고 Y 를 결과로 가지는 등식이 그대로 남아있다면 이것은 순환적이라는 지적이다.

둘째, 필은 간섭 개념을 제시한 후 X_i 가 x_i 를 가지게 될 때, X_i 의

13) Pearl (2000), p. 70.

14) Woodward (2001), p. 8.

Y 에 대한 인과적 영향력(causal effect)을 $Pr(y | do x_i)$ 에 의해 계산한다. 우리의 직관에 따르면 인과적 영향력은 간섭에 의한 변화만을 나타내야 한다. 하지만 펄이 제시한 간섭 개념에 따르면 간섭에 따른 인과적 영향력이 그렇게 계산되지 않는다. 즉 펄의 간섭 개념에 따르면, $Y = aX_i + U$ 라는 등식에서 $do x_i$ 라는 간섭이 있었다면 Y 에 대한 그 인과적 영향력은 $Pr(ax_i)$ 가 아니라 $Pr(ax_i + U)$ 라고 계산되고, 이것은 간섭에 따른 변화만을 계산하지 않고 Y 에 영향을 끼치는 다른 원인들의 영향력까지 함께 계산하는 것이다. 우드워드도 그와 같은 방식의 계산법이 인과적 영향력을 정확하게 측정하지 못한다고 지적한다.

두 번째 문제에 대한 우드워드의 입장은 $Pr_{do x_i, B_i}(Y) - Pr_{do x^*, B_i}(Y)$ 에 의해 인과적 영향력을 계산하는 것이다.¹⁵⁾ 우드워드는 펄과 달리 X_i 의 Y 에 대한 인과적 영향력을 X_i 값의 변화에 따른 Y 값의 변화에 의해 계산해야 한다고 보고, 이 때 인과적 영향력은 배경 맥락 B_i 에 상대적일뿐 간섭하는 대상과 그 결과 사이의 관계를 언급하지 말아야 한다고 주장한다. 이와 같은 인과적 영향력의 계산이 가능하기 위해 우드워드는 간섭 개념이 간섭하는 대상만이 아니라 두 번째 변수와 연관하여 제시되어야 한다고 주장한다. 우드워드는 이 점을 간섭의 상대성(relativity)이라고 제시한다. 즉 간섭은 어떤 변수 X_i 에 대한 간섭이 아니라 Y 와 관련된 X_i 에 대한 간섭이라는 것이다.

간섭의 상대성은 첫 번째 문제에 대해 우드워드가 해결책을 제시하는 데에도 도움이 된다. X_i 를 원인으로 하고 Y 를 결과로 가지는 등식을 그대로 두면 X_i 가 Y 의 원인인지를 따지기 위해 X_i 에 간섭을 하는 작업이 아무런 의미를 찾을 수 없고, 순환적이라는 비판을 받게 된다. 따라서 X_i 가 결과로 나타나는 등식을 제외하고 나

15) Woodward (2001), p. 8.

머지 등식은 그대로 두라는 펴의 규정은 그 나머지 등식 중에 X_i 를 원인으로 하고 Y 를 결과로 가지는 등식까지도 제외하는 좀 더 엄밀한 제한이 필요하다. 우드워드스는 간섭의 정의를 제시할 때 두 번째 변수를 언급할 필요가 있다고 주장하지만, 그 간섭의 상대성을 첫 번째 문제에 대한 해결 방안으로 제시한 것인지는 불분명하다. 그렇지만 Y 와 관련된 X_i 에 대한 간섭이기 때문에 X_i 를 원인으로 하고 Y 를 결과로 가지는 등식도 X_i 가 결과로 나타나는 등식과 함께 간섭이 일어날 때 제외되어야 한다고 보는 것이 우드워드스의 지적이다.

우드워드스는 펴가 제시한 원자적 간섭 개념이 조종을 단지 계산을 하기 위한 도구적 장치라고 평가한다.¹⁶⁾ 펴는 다른 인과 관계를 변화시키지 않고 문제의 변수 값을 조종할 수 있는 방안을 제시하는 데에 초점을 맞추고 있다는 것이다. 우드워드스는 이러한 펴의 접근 방식이 잘못되었다고 비판하지는 않지만, 간섭의 의미가 목적에 따라 다르게 구체화될 수 있다고 주장한다. 우드워드스는 인과의 본성을 탐구하기 위해 제시하는 자신의 간섭 개념은 펴가 제시하는 간섭 개념과 다르다고 주장하며, 그 이유로 간섭 개념을 제시하는 목적의 차이를 제시한다.¹⁷⁾ 즉 우드워드스는 자신의 간섭 개념 규정이 인과 관계의 진리 조건을 제시하려는 의미론적 작업으로, 통계적 정보로부터 인과 관계를 추론해내기 위해 필요한 조건을 규정하려는 펴의 작업과는 다르다고 주장한다. 과연 이 두 간섭 개념이 다른 것인지의 문제는 이후 논의하기로 하고, 일단 우드워드스의 간섭 개념을 좀 더 명확하게 이해해 보기로 한다.

우드워드스는 다음과 같은 점을 유의하며 자신의 간섭 개념을 규정한다.¹⁸⁾ 첫째, 간섭받는 것의 값은 전적으로 실험자의 간섭에 의

16) Woodward (2003), p. 110.

17) Woodward (2003), p. 110.

18) Woodward (2003), pp. 95-98.

해 결정된다. 둘째, 실험자의 간섭이 검사하려는 인과 관계의 결과와 관련을 지니는 어떤 다른 원인, 즉 검사하려는 인과 관계의 원인이 아닌 다른 원인과 상관 관계를 가진다면 실험의 신뢰도는 낮아질 것이다. 셋째, 간섭이 검사하려는 인과 관계의 결과 부분에 있는 것을 다른 경로를 통해 일으켜도 안 된다는 것이다.

이상의 특징을 반영하는 간섭 개념을 규정하기 위해 우드워드스는 Y 와 관련된 X 에 대한 간섭 변수(intervention variable) IV 를 정의하고, IV 에 근거해 간섭 IN 을 정의한다.¹⁹⁾ 먼저 IV 는 다음과 같다.²⁰⁾

19) 우드워드스는 이전에 WIN 과 SIN 등으로 간섭 개념을 정의하였으나 (1997, 2000) 이후 본 논문에서 논의하고 있는 것처럼 IV 와 IN 을 새롭게 도입하여 간섭 개념을 정의한다 (2003). 우드워드스가 IV 와 IN 을 통해 제시한 간섭 개념은 이전의 SIN 에 해당한다.

20) Woodward (2003), p. 98. 우드워드스는 $I1$ 에서 $I4$ 까지의 IV 정의가 불충분하다는 점을 지적하고 이를 보완한다. 먼저 문제점을 살펴보자. C 가 X 와 Y 의 공통 원인이고, X 와 Y 사이에는 다른 인과 관계를 가지지 않는 다음의 등식 체계를 가정해보자.

$$X=aC$$

$$Y=bC$$

이때 Y 와 관련된 X 에 대한 조종 M 이 b 의 값을 변화시킨다고 해보자. M 은 C 와 Y 사이의 관계를 변화시키는 것이다. 그렇다면 M 은 Y 와 관련된 X 에 대한 간섭이라고 할 수 있는가? 간섭이라고 한다면 위의 등식 체계가 보여주는 바와 달리 X 가 Y 의 원인이 된다. 이 문제를 해소하기 위해 우드워드스는 M 을 간섭이라고 부르지 않을 방안을 모색한다.

우드워드스는 두 가지 해결책을 제시한다. 첫째는 M 이 $I3$ 을 위반한다고 판단하는 것이다. M 은 X 에 대한 간섭이지만 X 를 통하지 않고 Y 의 값도 변화시킨다고 보는 것이다. M 은 C 와 Y 사이의 관계를 변화시키기 때문에 C 값이 고정되어 있어도 Y 값이 달라지도록 한다. 그것은 M 이 Y 값에 직접 영향을 끼치는 것이라고 할 수 있고, 따라서 M 은 $I3$ 의 위반으로 간섭이 되지 못한다. 둘째 해결책은 C 와 Y 의 관계를 변화시키는 M 이 간섭일 수 없다고 규정하는 새로운 조건 $I5$ 를 덧붙이는 것이다.

$I5$. I 는 X 와 Y 의 관계 밖에 있는 Y 의 원인 Z 와 Y 사이의 관계를 변화시키지 않는다.

IV

- I1. I 는 X 의 원인이다.
- I2. I 는 X 의 원인인 모든 다른 변수에 대해서 스위치 역할을 한다. 즉 I 가 특정한 값을 가지는 경우 X 는 더 이상 다른 변수의 값에 의존하지 않고 I 가 가진 값에만 의존한다.
- I3. I 로부터 Y 로 쓰여진 모든 인과 관계의 화살표는 X 를 통하여 진행된다. 즉 I 는 Y 를 향하여 직접 인과의 화살표를 쓰지 않는다. 또 I 는 X 와 구분되는, Y 의 원인이 되는 다른 어떤 것의 원인이 아니다.
- I4. I 는 X 를 통하지 않고 Y 의 원인이 되는 변수 Z 와 확률적으로 독립적이다.

이제 우드워드는 간섭 IN을 IV에 근거하여 다음처럼 정의한다.²¹⁾

IN

I 가 어떤 값 $I = z_i$ 를 가지는 것은 Y 와 관련된 X 에 대한 간섭이다 iff I 가 Y 와 관련된 X 에 대한 간섭 변수이고 $I = z_i$ 는 X 가 어떤 값을 가지는 실제 원인이다.

5. 두 가지 간섭 개념?

우드워드의 간섭 개념은 펄의 간섭 개념과 명확히 구분되는가? 우드워드는 자신의 간섭 개념이 의미론적 작업인 반면 펄의 간섭 개념은 통계적 정보로부터 인과 관계를 추론하는 것을 목적으로 하는 방법론적 작업이기 때문에 서로 다르다고 주장한다. 간섭의 의미가 목적에 따라 다르게 구체화될 수 있다는 주장이다. 우드워드의 주장을 검토해보자.

우드워드는 펄의 간섭 개념이 인과적 마코프 조건(causal Mark

두 해결책 모두 IV 정의를 보완하는 데에 충분하다.

21) Woodward (2003), pp. 98-99.

ov condition)을 전제하는 반면 자신의 간섭 개념은 그렇지 않다고 주장한다.²²⁾ 분명히 인과적 마코프 조건을 전제하는 것과 전제하지 않는 것은 커다란 차이점이다. 하지만 우드워드도 지적하는 것처럼 인과적 마코프 조건을 전제하는 것이 단지 조종과 간섭의 결과를 추론하고 계산하기에 유용하기 때문이라면, 인과적 마코프 조건을 전제하지 않는다는 점만으로 두 사람의 간섭 개념이 근본적으로 다르다고 말하기 어렵다. 우드워드의 간섭 개념에 인과적 마코프 조건을 덧붙이는 것이 무모순적이고, 나아가 인과적 마코프 조건을 전제하지 않는 이유가 단지 인과의 여러 개념을 포섭하기에 용이할 듯 하다는 가정만이라면 두 사람의 간섭 개념이 다르다고 판단하는 것은 성급할 수 있다는 것이다.

우드워드는 펄의 간섭 개념이 자신의 I3과 I4에 해당하는 조건을 제시하지 않는다고 지적하며 자신의 간섭 개념이 펄의 간섭 개념과 다를 뿐만 아니라 펄의 간섭 개념보다 더 구체적이라고 주장한다.²³⁾ 펄의 간섭 개념에 I3과 I4에 해당하는 조건이 없다는 우드워드의 지적은 옳다. 하지만 그렇다고 해서 우드워드의 간섭 개념이 펄의 간섭 개념과 다르다고 판단하는 것은 성급할 수 있다. 왜냐하면 같은 간섭 개념이라도 제시하는 방식이 다를 수 있기 때문이다. 따라서 한 번 더 생각해봐야 하는 문제는 I3과 I4에 해당하는 조건이 없어도 우드워드도 I3과 I4에 의해 만족시키고자 하는 점을 펄의 간섭 개념이 충족하고 있는가 하는 것이다.

펄의 간섭 개념에 따르면, 간섭되는 것의 값을 상수 x_i 로 정하는 등식 $X_i = x_i$ 를 새롭게 집어넣는 것이 간섭이다. 이렇게 새로운 등식을 넣는 대신 X_i 를 좌변에 가지는 기존의 등식은 제외하게 된다. 새로 들어가는 등식에는 Y 나 Z 등의 다른 변수가 전혀 나타나지

22) Woodward (2001), p. 108.

23) Woodward (2003), p. 110.

않는다. 따라서 그런 다른 변수의 발생을 우려하는 I3과 I4는 필의 간섭 개념 정의에서 전혀 고려하지 않아도 될 부분이다. 물론 필은 그런 조건을 첨가할 수도 있다. 즉 간섭을 통해 새로 들어가는 등식에는 다른 변수가 나타나지 말아야 한다는 조건을 덧붙일 수 있다. 그러나 그러한 덧붙임은 필의 간섭 개념을 다른 간섭 개념으로 만든다고 하기 어렵다. 우드워드도 필의 간섭 개념이 암묵적으로 전제하고 있는 바를 구체적으로 제시하고 있을 뿐이다. 그러나 그 점 때문에 우드워드의 간섭 개념이 필의 간섭 개념과 다르다고 결론지을 수는 없다. 우드워드가 제시한 I1에서 I4까지의 조건에서 우드워드의 간섭 개념과 필의 간섭 개념 사이의 근본적 차이점을 찾기는 어렵다.

우드워드가 두 간섭 개념의 차이점으로 지적하고 있는 또 다른 측면은 필의 간섭 개념이 인과 관계에 근거하고 있고, 그에 따라 인과를 이해하고자 하는 작업에는 순환성의 문제 때문에 부적합한 반면 자신의 간섭 개념은 그렇지 않다는 것이다.²⁴⁾ 우드워드는 이 주장의 근거로 염두에 두고 있는 것은 아마도 간섭의 상대성일 것이다. 간섭의 상대성이란 간섭을 정의할 때 두 번째 변수를 언급할 필요가 있다는 것이다. 즉 간섭은 언제나 다른 변수에 상대적으로 어떤 변수에 대해 일어나는 상대성을 지니며, 단순히 X 에 대한 간섭이 아니라 Y 와 관련된 X 에 대한 간섭이다.

우드워드가 지적하는 것처럼 필은 간섭 개념을 정의할 때 두 번째 변수를 언급할 필요성을 지적하지 않는다. 그리고 우드워드의 지적처럼, X 를 원인으로 하고 Y 를 결과로 가지는 등식을 그대로 두면 X 가 Y 의 원인인지를 따지기 위해 X 에 간섭을 하는 작업은 약순환일 수 밖에 없다. 따라서 X 가 결과로 나타나는 등식을 제외하고 나머지 등식은 그대로 두라는 필의 규정은 그 나머지 등식

24) Woodward (2003), p. 110.

중에 X 를 원인으로 하고 Y 를 결과로 가지는 등식까지도 제외하는 좀 더 엄밀한 제한을 필요로 한다. Y 와 관련된 X 에 대한 간섭이기 때문에 X 를 원인으로 하고 Y 를 결과로 가지는 등식도 X 가 결과로 나타나는 등식과 함께 간섭의 영향력을 올바르게 측정하기 위해서는 제외해야 한다.

간섭 개념에 대한 이러한 제한을 가능하게 한 것이 우드워드가 제시한 간섭의 상대성이다.²⁵⁾ 반면 펄의 간섭 개념 정의에는 간섭의 상대성에 대한 언급이 없다. 그러나 간섭의 상대성에 대한 언급이 없다는 이유만으로 우드워드의 간섭 개념이 펄의 간섭 개념과 근본적으로 다르다고 판단하는 것은 성급할 수 있다. 적어도 펄의 간섭 개념 정의에 간섭의 상대성이란 특성을 추가하는 데에는 아무런 문제가 없기 때문이다. 나아가 펄의 의도한 간섭 개념은 간섭의 상대성으로 제시한 우드워드의 조건을 만족시킬 필요가 있는 것이다. 따라서 펄의 간섭 개념이 우드워드의 간섭 개념과 근본적으로 다르다고 말하기 보다는 펄이 분명히 제시하지 못한 부분을 우드워드가 명료하게 만들었다고 말하는 것이 적절하다.

마지막으로 우드워드는 자신의 간섭 개념이 비의인화(non-anthropomorphism)의 특성을 지닌다고 주장한다.²⁶⁾ 즉 자신의 간섭 개념은 인간의 행위에 따른 것이라는 전제를 지니지 않는다는 것이다. 하지만 펄의 간섭 개념 역시 간섭이 인간의 행위에 따른 것이라는 전제를 두지 않는다. 따라서 이 점 역시 펄의 간섭 개념이 암묵적으로 전제하고 있는 바를 구체적으로 명시한 것에 불과하고, 간섭 개념과 근본적으로 다르다고 말하기 어렵다.

이처럼 펄과 우드워드의 간섭 개념 사이의 차이점은 근본적인 것이라고 보기 어렵다. 물론 간섭의 상대성 등의 특성을 밝혀낸 우

25) Woodward (2001), p. 103.

26) Woodward (2001), p. 103.

드워드¹⁾의 지적은 사소한 지적이 아니다. 그러나 간섭의 상대성은 펄의 간섭 개념이 보완할 수 있는 측면이다. 그 차이점은 두 간섭 개념이 근본적으로 다른 개념이라고 결론짓는 근거가 되기에는 불충분하다. 이처럼 두 간섭 개념 사이에 근본적 차이점을 찾기 어렵다면, 목적의 차이가 서로 다른 간섭 개념을 만든다는 우드워드의 주장 역시 적절한 근거를 찾기 어렵다. 목적의 차이가 간섭 개념을 서로 다른 방식으로 제시하도록 할 수는 있을지 몰라도 두 서로 다른 개념을 만들지는 않을 것이라는 것이 설득력을 지닌 직관일 것이다.

6. 나오기

과학자들이 인과를 어떻게 이해하고 사용하고 있는지를 분석하는 작업은 조종가능성 이론이 비환원주의적 접근 방식이기 때문에 가지는 한계가 있다고 하더라도 인과의 본성을 밝히려는 시도에도 중요한 역할을 할 수 있다. 비록 순환성의 문제 등이 조종가능성 이론의 심각한 난점으로 제기되고 있지만, 조종의 의미를 간섭 등에 초점을 맞춰 정교화하려는 노력은 인과의 이해를 높일 수 있는 방안을 제시한다. 행위자의 행위에 초점을 맞추며 환원주의적 접근 방식을 고수하기보다는 순환성을 인정하되 조종가능성으로의 이해가 인과의 본성을 더 쉽게 이해시켜 줄 수 있다는 점을 강조하는 접근 방식이 철학적으로도 평가절하될 필요는 없다. 목표는 인과의 이해이기 때문이다.

본 논문은 특히 펄이 제시한 원자적 간섭 개념과 우드워드가 제시한 간섭 개념을 비교하며, 둘 사이의 차이가 근본적인가를 검토했다. 우드워드는 펄의 원자적 간섭 개념이 계산을 하기 위한 도구

적 장치일 뿐이라고 보고, 인과의 본성을 탐구하기 위해 제시한 자신의 간섭 개념은 펄의 간섭 개념과 다르다고 주장하지만, 본 논문은 그와 같은 우드워드 주장이 충분한 근거를 지니지 못한다고 논증했다. 이 과정에서 논자는 펄과 우드워드 모두 인과의 조종가능성 이론을 발전시키기 위해 간섭 개념을 정교화하려고 시도하였고, 등식과 방향 그래프를 사용하여 인과 관계를 분석하려 했다는 점에 주목한다. 이는 인과의 조종가능성 이론이 어떤 방향으로 발전될 수 있는가를 전망할 수 있도록 한다. 그와 같은 전망을 좇아 인과의 의미가 사회과학에서 어떻게 구체화되고 있는지를 밝히는 작업은 논자의 추후 작업 목표가 될 것이다.

참고 문헌

- 이영의(2005), “인과적 마코프 조건과 비결정론적 세계”, 『논리연구』 제8집 1호 pp. 47-67.
- Collins, J. & Hall, N. & Paul, L. A.(2004), *Causation and Counterfactuals*. MIT Press.
- Glymour, C.(2004), “Critical Notice: James Woodward”, BJPS 55 pp. 779-790.
- Hausman, D.(1998), *Causal Asymmetry*, Cambridge Univ. Press.
- Menzies, P. & Price, H.(1993), “Causation as a Secondary Quality”, BJPS 44, pp. 187-203.
- Pearl, J.(2000), *Causality*, Cambridge Univ. Press.
- Price, H.(1991), “Agency and Probabilistic Causality”, BJPS 42, pp. 157-176.
- Spirtes, p. & Glymour, C. & Scheines, R.(2000), *Causation, Prediction, and Search*. 2nd edition. MIT Press.
- von Wright, G. H.(1993), “On the Logic and Epistemology of the Causal Relation” in Sosa, E. & Tooley, M., *Causation*. Oxford Univ. Press.
- Woodward, J., “Causation and Manipulability”, The Stanford Encyclopedia of Philosophy(Fall 2001 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2001/entries/causation-mani/>>.
- _____ (2003), *Making Things Happen*. Oxford Univ. Press.

동덕여자대학교 교양학부

Email: yyeo4@hanmail.net

The Manipulability Theory of Causation and Intervention

Yeongseo Yeo

Manipulability theories of causation have considerable intuitive appeal and are popular among scientists. The central idea of manipulability theories of causation is that causal relationships are relationships that are potentially exploitable for purposes of manipulation and control. That is, if C is a cause of E , then if someone can manipulate C in the right way, this should be a way of manipulating or changing E . In order to clarify the central concept of this theory, manipulation, recent developments in manipulability theories of causation concentrate on the concept 'intervention'. Pearl suggests atomic intervention, which brings about a surgical change, using equations and directed acyclic graph. It controls the variable in question leaving aside the rest of the causal relations. Woodward claims that Pearl's characterization of intervention is merely calculational different from his own which aims to clarify the meaning of causation. In this paper, I argue that Woodward's characterization of intervention is not different from Pearl's. I show that Woodward's characterization of intervention materializes Pearl's concept of intervention.

【Key Words】 causation, manipulation, intervention, Pearl,
Woodward