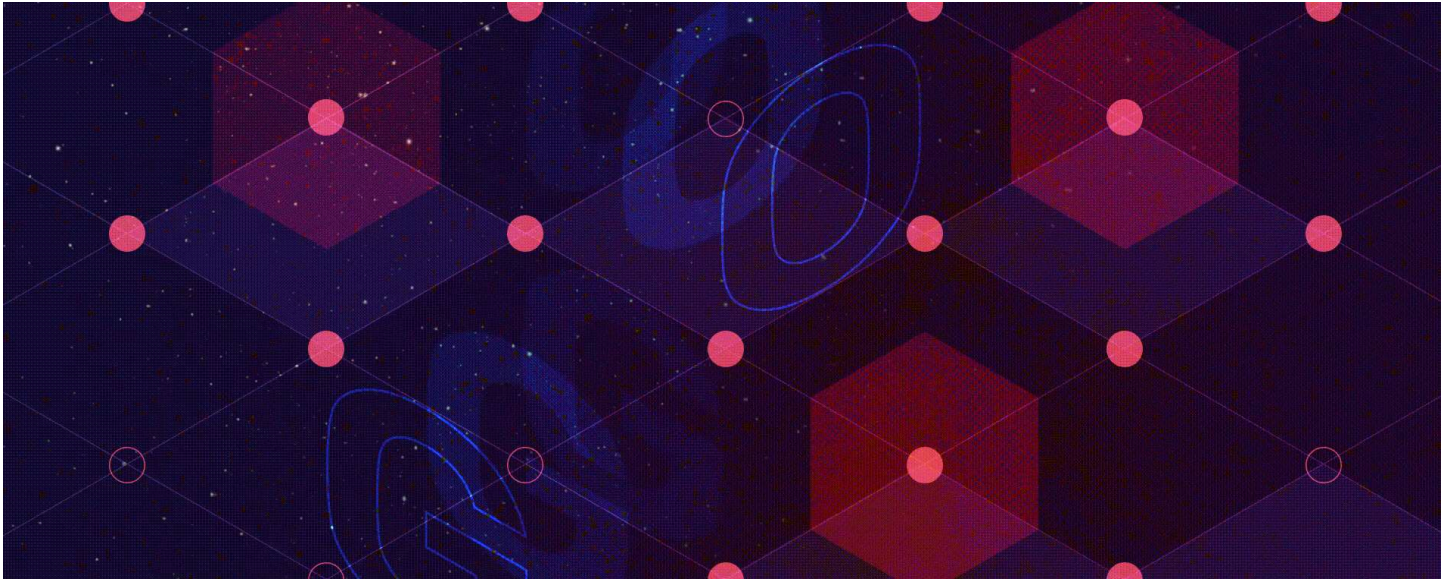


네트워크의 창발

2019년 9월 17일

손승우



2016년 3월 9일 세계의 이목이 구글 딥마인드 챌린지 매치가 열리는 서울 포시즌스 호텔에 집중되었다. 열아홉 줄 바둑판 위에서 펼쳐진 인공지능과 인간의 역사적인 대결은 유튜브를 통하여 전 세계로 생중계되었다. 일주일 동안 펼쳐진 다섯 번의 대결에서 결국은 4승 1패로 인공지능이 승리했다.

'돌부처' 이창호 9단의 전성기보다 낫았다고 평가되는 치밀하고 정교한 수읽기와 정확한 집계산 능력으로 알파고가 첫 세 대국을 연달아 승리하며 5번기의 우승을 확정 짓는 순간... 인공지능의 무시무시한 강력함과 인간 최고 바둑기사인 이세돌 9단의 떨리는 손끝을 목도하며 절망감을 느꼈다. 이어진 3월 13일의 제4국에서 '신의 한 수'라 평가되는 백 78 수를 통해 이세돌 9단이 1승을 거두었을 때는 인류의 희망을 본 듯이 기뻐다. 다시 이틀 뒤 이어진 최종국의 패배는 후속편이 준비되어 있을 것 같은 SF 영화가 끝난 뒤 발걸음이 떨어지지 않는 영화관 밖을 나서야 할 때처럼 아쉬웠다.



그림1

[Nature](#), Google Deepmind

흑 소목, 백 화점, 흑 소목, 백의 양화점... 대략 서른 개의 바둑돌이 놓일 때까지 아득하던 연결은 50수를 지나고, 100수를 지나며 점점 뚜렷해진다. 견고한 연결과 너른 영역을 모두 염두에 두고 치열한 두뇌 싸움을 펼치는 바둑에서는 단순하게만 따져도 $(19 \times 19)!$ 개의 가능한 경우의 수가 있는데, 이는 1뒤에 0이 845개나 쓰인 큰 수이다. 바둑의 규칙에 따라 놓을 수 없는 곳을 제하면 이를 좀 많이 줄일 수 있어 0을 170개까지 줄일 수 있다.[1] 그래도, 아직도 소위 '에딩턴 수 Eddington number'라고 알려진, 관측 가능한 우주에 존재하는 모든 양성자의 수로 추정되는 10의 80승[2]보다도 동그라미가 90개 많다. 이런 우주보다도 '더 넓은 우주'를 어렵פות이나마 탐색하는 인간의 두뇌와 이를 압도하는 인공지능이 대단하게 느껴진다.

통계물리학에는 바둑과 닮은, 하지만 좀 더 간단한 문제가 있다. 다시 열아홉 줄의 바둑판으로 돌아가 보자. 이번엔 흑 돌만 쥐고 첫 수가 꼭 소목이 아니어도 좋으니 마구잡이로 몇 개의 돌을 가로세로 열아홉 줄의 교차점에 놓아본다. 시작한 김에 서른 개쯤 놓아 보자. 어떤 연결이 있을까? 몇 개의 흑돌은 나란히 놓여 연결되었다. 두 개가 연결된 것도 있고, 세 개가 연결된 것도 있다. 이렇게 100개의 돌을 놓아 보면, 어떤 돌은 벌써 변에 놓이게 된다. 제법 집 같이 연결된 곳도 보일 것이다. 하지만 아직 빈 곳이 더 많다. 크고 작게 연결된 돌들이 있다.

상하좌우로 몇 개의 돌이 연결된 송이 구조를 '무리 cluster'라 한다. 다양한 크기의 무리가 존재한다. 이렇게 몇 개의 돌을 더 놓다 보면 한 '아주 커다란 무리 giant cluster'가 상변과 하변을 연결하거나, 좌우를 연결할 수 있을 것이다. 이렇게 양끝을 연결하는 무리를 '연결무리 spanning cluster'라고 한다. 연결무리로 인해서 바둑판의 한쪽 끝이 다른 쪽 끝에 닿는다. 바둑판을 가로지르는 꾸불꾸불하지만 기다란 줄이 하나 생긴 것이다.

‘과연 몇 개의 돌을 놓아야만 이렇게 커다란 무리가 상하 혹은 좌우를 연결할까?’, ‘이때 각 크기의 무리들이 몇 개나 있을까?’, ‘그 무리의 크기 분포는 수월히 계산해 볼 수 있을까?’



그림2 해머슬리 John Michael Hammersley

이와 같은 수학적 모델을 도입한 것은 1957년 영국의 수학자 해머슬리¹ John Michael Hammersley, 1920~2004와 서른 살의 젊은 통계학자 브로드벤트² Simon Ralph Broadbent, 1928~2002였다.[3] 물론 이들이 바둑판에서 바둑돌의 연결을 생각했던 것은 아니고, 과수원에서 작물의 전염병이 어떻게 퍼질지와 탄소 알갱이가 가득한 가스 마스크에서 공기의 여과에 관하여 연구하던 차였다.[4] 그래서 이를 ‘스미기 모형’ 혹은 ‘투과 모형’, ‘여과 모형’ percolation model이라고 부른다. 흑돌의 연결무리가 양끝을 이은 바둑판에는 먹과 같은 검은 줄이 스며있고, 그 길을 따라 물이 ‘투과’되어 흐르는 듯하다. 바둑판이 퍼콜레이트(!)된 것이다.

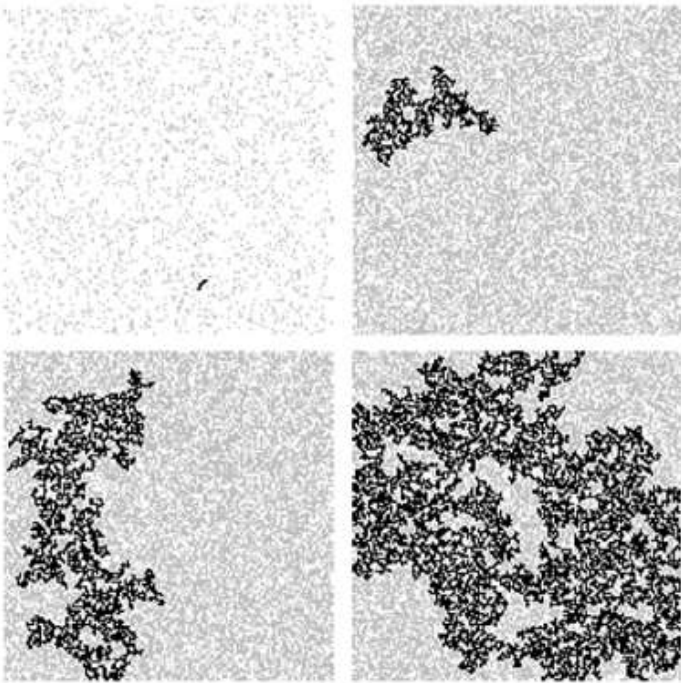
이렇듯 여과 모형은 과수원 작물에 퍼지는 감염병에 대해 설명한다. 양끝을 연결한 연결무리는 한 과수원에서 다른 과수원으로 감염병의 전염을 뜻한다. 매년 산림청을 긴장하게 하는 솔잎혹파리나 소나무재선충병의 확산과 산불이 퍼져 나가는 것, 지리산에 방사된 반달가슴곰이 백두대간을 따라 이동하는 것도 서로 다르지 않다. 이차원 평면에서의 여과 이론³ percolation theory[5]을 삼차원 공간으로 확장해 보면, 가스 마스크로 여과되는 공기를 설명하는 것은 암반을 따라 흐르는 지하수를 설명하는 것과 닮아있다. 암반 사이에 스며있는 석유가 땅속에 어떤 규모로 분포하는가와 같은 문제가 되는 것이다.

¹ <어벤저스> 시리즈에서와 같은 무한의 힘은 없겠지만 그 테서랙트가 맞다. 차라리 크리스토퍼 놀란 감독이 영화 <인터스텔라>에서 책장으로 표현한 4차원의 시공간 테서랙트가 좀 더 가깝겠지만.

² N 개의 점 사이에 가능한 연결선의 수 $M = \frac{N(N-1)}{2}$.

이를 3차원에서 그치지 않고 4차원, 5차원으로 확장할 수 있을까? 4차원 공간을 당장 머릿속에 그리는 것은 어려우나, 바둑판과 같은 사각격자를 확장하여 각설탕 같은 정육면체^{cube}로 가득 찬 단순입방격자^{simple cubic lattice}를 만드는 것처럼, 테서랙트^{tesseract}라 불리는 4차원 초입방체^{hypercube}를 잘 쌓으면 된다.¹ 2차원 평면의 사각격자에서는 상하좌우 네 개의 이웃이 있듯이, 3차원 단순입방격자에는 상하좌우앞뒤 여섯 개의 이웃이 있다. 4차원 초입방격자에는 몇 개의 이웃이 있을까? 바로 여덟이다. 이를 프로그램 언어로 표현하는 것은 사실 좀 더 쉽다. 더 이상 공간적인 구조의 상상은 필요 없고, 수학적 규칙성만을 따져서 배열^{array}이라는 자료구조를 잘 활용하여 네 개의 번호를 갖는 4차원 배열을 만들면 된다. 이런 추상화가 가능한 것이 수학의 아름다움일 것이다.

이러한 상상을 계속 확장해나가면 어떻게 될까? 차원이 더해질수록 방 안의 더 많은 이웃을 만날 수 있다. 바둑판의 교차점은 2차원 평면이기에 네 개의 이웃밖에 없었다면, 3차원에서는 여섯 개, 4차원에서는 여덟 개의 이웃을 가질 수 있다. 이를 계속 늘려 가면, 한 공간 안에 아직 연결되지 않은 N 개의 점이 모두 연결 가능한 잠재적 이웃인 상태를 상상할 수 있다. 우리가 한 개의 바둑돌을 놓아 한 점을 차지했듯이, 임의로 선택한 아직 서로 연결되지 않은 두 개의 점을 이으며 하나의 연결선을 추가할 수 있다. 이를 계속 반복해 나가다 보면 적당히 많은 점이 연결되는 때가 있을 것이다.



그림³ 바둑판과 같은 2차원 사각격자에서 점의 개수를 늘려가는 site percolation의 예. 차례대로 10%, 55%, 58%, 59.2746%가 채워진 상태로 진한 검은 색 무리가 가장 큰 무리이다.

K. Christensen and R. Moloney, "Complexity and Criticality" (2005).

10개의 점을 예로 들어보자. 과연 몇 개의 연결선으로 연결이 가능할까? 10개의 점을 줄줄이 엮자면 9개의 연결선이면 충분하고, 별 모양처럼 가운데 하나를 두고 나머지 9개의 점을 모두 연결하는 데에도 9개의 연결선이 필요하다. 하지만, 이제 마구잡이로 두 점을 연결한다는 데에 주목하자. 앞선 예에서는 하나의 연결도 낭비 없이 하나의 점을 더 연결하는 데 쓰였지만, 잘못하여 삼각형 모양이나 사각형 모양의 고리를 만드는 순간 하나의 연결선은 추가적인 점을 연결하지 못하고 낭비될 수밖에 없다. 3개의 점은 그사이에 가능한 연결선이 3개뿐이지만, 4개의 점은 연결선 6개, 5개의 점은 10개로 증가하여 10개의 점 사이에는 45개의 가능한 연결선이 존재한다.² 그중에 9개는 1/5로 가능한 연결선 중 20%만으로 아주 효율적으로 연결한 예이다.

바둑판 모양의 사각격자에서 경계를 잇는 연결무리가 361(=19×19)개의 모든 점을 연결하지는 않듯이, 여기서도 적당히 많은 수의 점이 아주 커다란 무리를 이루면 족하다. N개의 점이 있을 때, 가장 커다란 무리가 N에 비견될 만한 크기를 이룰 때 '여과'가 이루어졌다고 본다. 그렇다면 N이 아주 커서 점이 아주 많다면, 과연 몇 개의 연결선이 연결될 때, 이 점들은 연결되어 퍼콜레이트(!)될까? 그 답은 흥미롭게도 ' $\frac{N}{2}$ '개의 연결선이면 충분하다'이다.[6] 이정도 연결선이면 가장 커다란 무리가 차지하는 상대적인 크기가 N이 아무리 커져도 작지만 0이 아닌 유한한 값을 갖기 시작한다. 이때가 바로 세상의 무수히 많은 점들이 연결되어 네트워크가 만들어지는 **네트워크 창발의 시점**이다.

³ (N - 1)개의 노드에 연결확률 p로 연결될 때, 각 노드의 연결선 수의 분포는 다음과 같은 이항분포를 따른다.

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-k-1},$$

여기서 $\binom{N-1}{k} = \frac{(N-1)!}{k!(N-k-1)!}$ 이다.

이렇게 고정된 N개의 점(노드node)이 마구잡이로 쌍을 이루어 M개의 연결선(링크link)으로 연결되어 있을 때, 각 노드는 연결확률 $p = \frac{2M}{N(N-1)}$ 를 갖고, 그 연결선 수degree는 종모양의 이항분포³를 따른다. 이를 에르되시Paul Erdős, 1913~1996와 레니Alfréd Rényi, 1921~1970의 '무작위 네트워크random network'라 한다. 이러한 복잡계 네트워크의 여과 이론은 우리에게 네트워크의 창발에 대한 이해뿐 아니라, 그 역과정인 네트워크의 붕괴가 어떠한 과정을 거칠 것인지에 대해서도 힌트를 준다. 교통망이나 전력망, 통신망, 인터넷 등이 사소한 오류로 망가지거나 계획적인 테러에 의해서 공격받을 때, 그 피해의 정도가 파국적인 블랙아웃blackout이 될지 국지적인 짧은 정전이 될지 가늠해 볼 수 있다.

올해 초 3월 미국에서 열린 미국물리학회APS March Meeting에서는 "네트워크 과학 20년: 구조에서 조절까지Twenty Years of Networks Science: From Structure to Control"라는 제목의 기조강연이 있었다. 강연자는 물론 복잡계 네트워크 이론의 선구자라 할 수 있는 바라바시Albert-László Barabási, 1967~였다. 최근의 네트워크 이론은 학문적 성숙기에 접어들어 다양한 분야에 응용되고 있다. 우리는 이제 소셜 미디어에서 가짜 뉴스의 확산이 우리 사회에서 전염병이 전파되는 것과 같고, 새로운 혁신적 아이디어의 전파와도 양상이 같다는 것을 알고 있다. 우리는 퍼콜레이션 이론을 이용하여 가짜뉴스와 질병의 확산을 억제하고, 새로운 아이디어의 확산은 더욱 활발히 퍼지도록 조절할 수 있다. 헤머슬리와 브로드벤트가 60년 전에 도입한 수학 이론이, 사회 연결망과 인터넷 연결 구조에 관한 실증적 데이터 연구로부터 시작된 지 이제 20년이 된 네트워크 과학[7]을 만나 새로운 지식을 생산한다.



그림 4-1 좌 에르되시Paul Erdős



그림 4-2 우 바라바시Albert-László Barabási

4-1 [Billy or Grace Tao](#), 4-2 [Wikipedia](#)

최근 인공지능경망의 연구는 빠르게 발전하고 있다. 신경과학 연구의 결과로 얻은 뇌 연결망의 이해는 새로운 인공지능경망 구조를 제안한다. 그중 순환 신경망recurrent neural network은 그 구조에 인간의 기억과 닮은 동적인 내부 상태를 가질 수 있고, 순차적으로 들어오는 입력값을 처리하는 데 유용하여 음성인식과 로봇 컨트롤 등에 활용된다. 최근 이를 축적 컴퓨팅reservoir computing 방법으로 학습시킨 인공지능경망은 예측불가능하다고 알려진 비선형 카오스 시스템의 예측에

도전하고 있다.[8] 그 인공지능경망의 한 가운데 레저버^{reservoir}에 또한 ‘무작위 네트워크’ 구조가 자리하고 있다. (Networks are Ubiquitous! 네트워크는 어디에나 있다.) 어떠한 네트워크 구조가 보다 나은 성능을 보일 것인가, 그런 네트워크 구조의 역할은 무엇인가에 대한 연구가 진행되고 있으며, 뛰어난 성능을 보이는 인공지능경망의 구조로부터 인간 뇌 신경망의 기능을 다시 이해하려는 연구도 활발히 진행 중이다. 마치 알파고의 수로부터 인류의 바둑에 대한 이해를 더할 수 있었던 것과 비슷하다.

어린 시절 인기를 끌었던 SF 영화 <터미네이터>의 후속편이 올해 개봉한다고 한다. ‘인류를 멸망시키려던 강인공지능 스카이넷은 어쩌면 누군가에 의해 설계된 것이 아닌 인공지능의 네트워크가 스스로 창발한 것은 아닐지?’ 아득한 상상을 하며 다시 나올 것 같지 않았던 SF 영화의 후속편을 기다려 본다.

참고문헌

1. Tromp and G. Farneback (2007), "Combinatorics of Go", In: van den Herik H.J., Ciancarini P., Donkers H.H.L.M. (eds.), "Computers and Games", CG 2006, Lecture Notes in Computer Science, vol 4630. Springer, Berlin, Heidelberg.
2. S. Eddington (1956), "The Constants of Nature", In J. R. Newman (ed.) The World of Mathematics², pp. 1074-1093, Simon & Schuster. https://en.wikipedia.org/wiki/Eddington_number
3. Bacaër (2011), "Percolation and epidemics (1957)", In: A Short History of Mathematical Population Dynamics, Springer, London.
4. 강병남 (2014), "폭발적 여과 전이 모형 및 이론", 학술원논문집(자연과학편) 제53집 2호 pp. 105-126.
5. Stauffer and A. Aharony (1994), "Introduction to Percolation Theory", 2nd ed., Taylor & Francis.
6. Erdős and A. Rényi (1959), "On Random Graphs. I", Publicationes Mathematicae, 6: 290-297.
7. 정하웅, 강병남 (2007), "복잡계 네트워크에 대한 최근 연구 동향", 물리학과 첨단기술 2007년 10월 제16권 10호; 강병남 외 (2004), "복잡계 네트워크의 구조적 특징과 동역학 현상", 새물리 48권 2호 pp. 115-141.
8. Pathak et al. (2018), "Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach", Physical Review Letters 120, 024102.