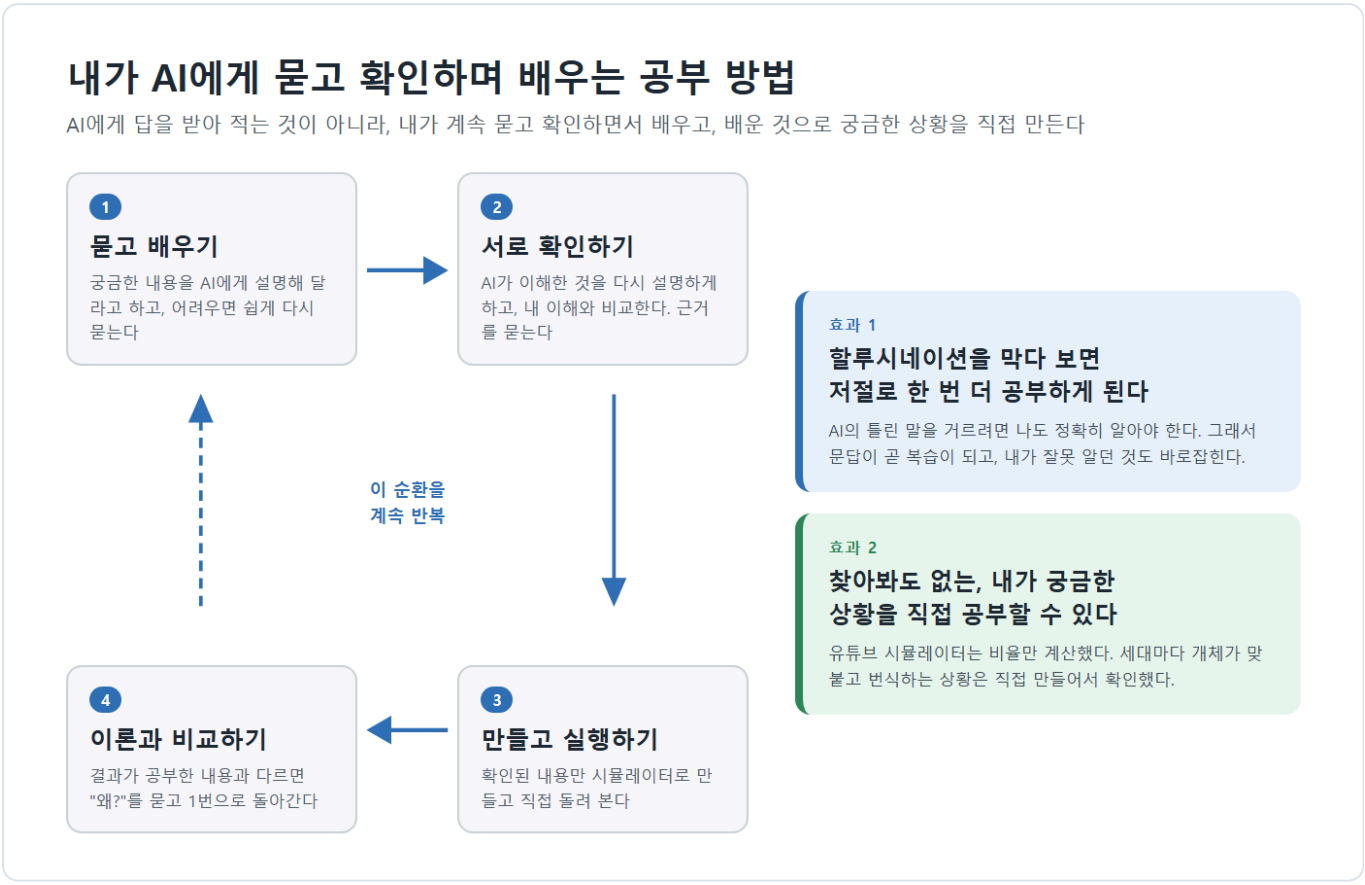
내가 AI에게 묻고 확인하며 배운 과정

저는『이기적 유전자』5장을 더 깊이 이해하고 싶어서 AI(Claude Code)에게 묻고 확인하며 공부했습니다. 아래는 제가 실제로 한 질문과, 그 질문으로 제가 확인하고 바꾼 것들입니다. 질문은 대화 기록에서 원문 그대로 옮겼습니다.



저는 묻고 배우기 → 서로 확인하기 → 만들고 실행하기 → 이론과 비교하기를 계속 반복했습니다.

1 묻고 배우기 — 이해될 때까지 다시 묻기

내가 한 질문

"이기적 유전자 5장 내용 설명해줘"
"비둘기파, 매파 게임이론을 상세히 설명해줘"

내가 한 일

책에서 읽은 ESS, 매파와 비둘기파의 점수표, 7대5 균형 계산을 AI의 설명과 하나씩 비교하며 다시 공부했습니다.

내가 한 질문

"너무 어려워. 쉽고 자세히 예를 들어 구체적으로 설명해줘"

내가 한 일

어려운 용어로 된 설명은 그냥 넘기지 않고, 숫자 예시와 비유로 다시 설명하게 했습니다. 그 뒤로는 계속 이 방식으로 설명을 요청했습니다.

> **내가 알게 된 것** — 모르면 넘어가지 말고, 내가 이해할 수 있는 말로 다시 물으면 된다.

2 서로 확인하기 — 내가 이해한 것이 맞는지 묻기

내가 한 질문

"집단 평균 먹이가 매파와 비둘기파 전체의 집단이야 아니면 같은 전략만 집단이야?"
"그러니까 비둘기파를 예로 들면 비둘기파 개체 중 1마리가 다음세대에 낳는 자식 수는 그 개체가 먹은 먹이 나누기 비둘기파 집단의 평균 먹이라는 거야?"

내가 한 일

제가 이해한 방식을 말로 정리해서 맞는지 물었습니다. 확인해 보니 나누는 기준은 같은 전략끼리의 평균이 아니라 **전체 집단의 평균**이었습니다. 같은 전략끼리 나누면 누구나 항상 본전이 되어 비율이 영원히 변하지 않는다는 것도 숫자로 확인했습니다.

> **내가 알게 된 것** — 제가 잘못 이해하고 있던 부분이었습니다. 확인하는 질문을 하지 않았다면 틀린 채로 넘어갔을 것입니다.

3 AI의 틀린 점 잡기 — 할루시네이션 거르기

내가 한 질문

"시간은 항목이 빠진거야? 시간도 넣었으면 해."

내가 한 일

AI가 세운 계획을 책 내용과 비교하다가, 책에 나오는 과시 비용(비둘기파끼리 서로 노려보며 시간을 쓰는 비용)이 빠진 것을 찾아냈습니다. 과시 비용을 넣게 했고, 그러자 균형 비율을 구하는 식이 바뀌었습니다.

> **내가 알게 된 것** — 내가 책 내용을 알고 있었기 때문에 AI가 빠뜨린 것을 찾을 수 있었다.

내가 한 질문

"보복자, 불량배, 시험보복자를 계산할 때는 결과로만 계산해줘. 예를 들어서 불량배와 불량배가 만나면 서로 공격하잖아. 그러니까 둘 다 도망가서 0\0이 되는거지"

내가 한 일

AI는 불량배끼리 만나면 동전 던지기로 승부를 정하도록 임의로 만들어 두었습니다. 저는 책의 불량배 설명대로라면 둘 다 도망가야 한다고 판단해서 고치게 했습니다. 4단계에 '둘 다 도망'을 넣자, 불량배끼리 있을 때 이상하게 점수가 높던 결과가 사라졌습니다.

> **내가 알게 된 것** — AI가 그럴듯하게 만든 부분도 책의 전략 설명과 맞는지 내가 직접 따져 봐야 한다.

4 근거 묻기 — "왜?"와 "그러면 무슨 문제가 생겨?"

내가 한 질문

"계산해서 마이너스가 나오면 어떻게 계산되는거야?"
"기본 점수가 의미하는게 뭐야? 기본 점수를 빼고 점수가 마이너스가 되면 0으로 처리하면 되지 않아? 이렇게 했을 때 생기는 문제가 뭐야?"

내가 한 일

제가 생각한 간단한 방법(마이너스는 0으로)에 어떤 문제가 있는지 근거를 물었습니다. 0으로 자르면 크게 다친 매파와 그냥 도망간 비둘기파가 똑같이 0점이 되어 매파가 실제보다 유리해지고, 실제로 돌려보니 매파가 99.9%까지 늘어났습니다(책의 균형은 58.33%). 그래서 모든 개체에게 기본 점수를 주는 방법을 택했습니다.

> **내가 알게 된 것** — 간단해 보이는 방법도 근거를 물어보면 숨은 문제가 드러난다.

5 내가 궁금한 상황 만들기 — 찾아봐도 없던 것

내가 한 질문

"내가 비슷한 실험에서는 점수를 먹이에 비유했는데 먹이를 1개 먹으면 생존 2개를 먹으면 생존 및 번식 0.5개 먹으면 50%로 생존. 이런 식으로 설명해줘"
"각각의 개체가 얻은 점수를 각자 계산해서 점수가 50이면 생존 100이면 자손 1마리 75면 생존 및 50퍼센트로 번식 ... 0이나 마이너스면 사망으로 해줘"

내가 규칙을 바꾼 이유

"이전 방식대로면은 원래 점수 50점이면 살 수 있는 상황에서 만약 다른 개체 수의 변화로 50점일 때 죽는 상황은 현실에서는 안 맞는 방식이잖아"

> **내가 알게 된 것** — 남이 만든 시뮬레이터는 정해진 방식만 보여 주지만, 직접 만들면 내가 현실에 더 가깝다고 생각하는 규칙으로 바꿔서 실험할 수 있다.

내가 한 질문

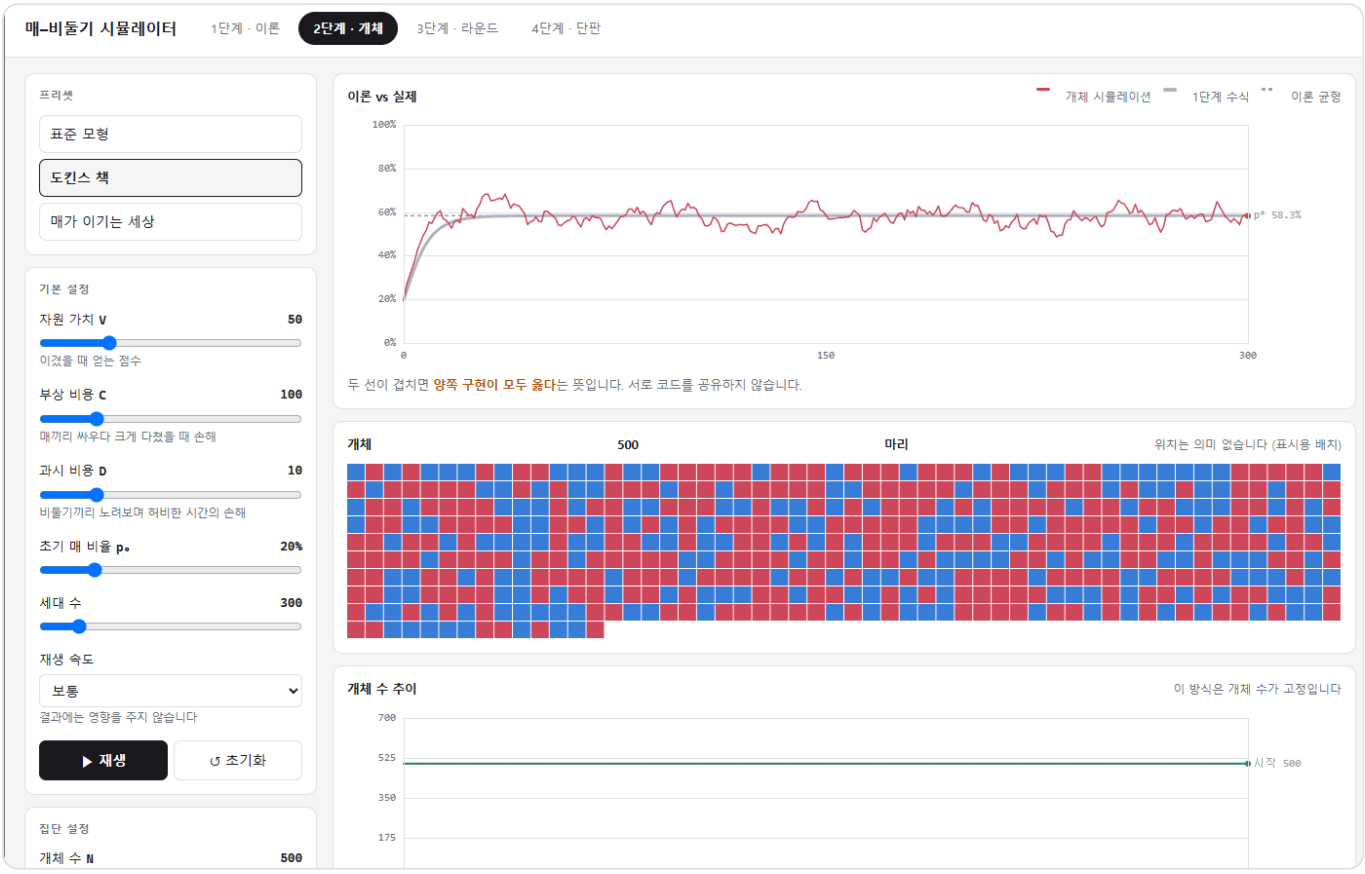
"각 개체별 대면이 실제로 마주치는 거야 아니면 확률로 계산하는 거야?"
"개체 수 추이가 항상 5000마리가 되는 이유가 뭐야?"

내가 한 일

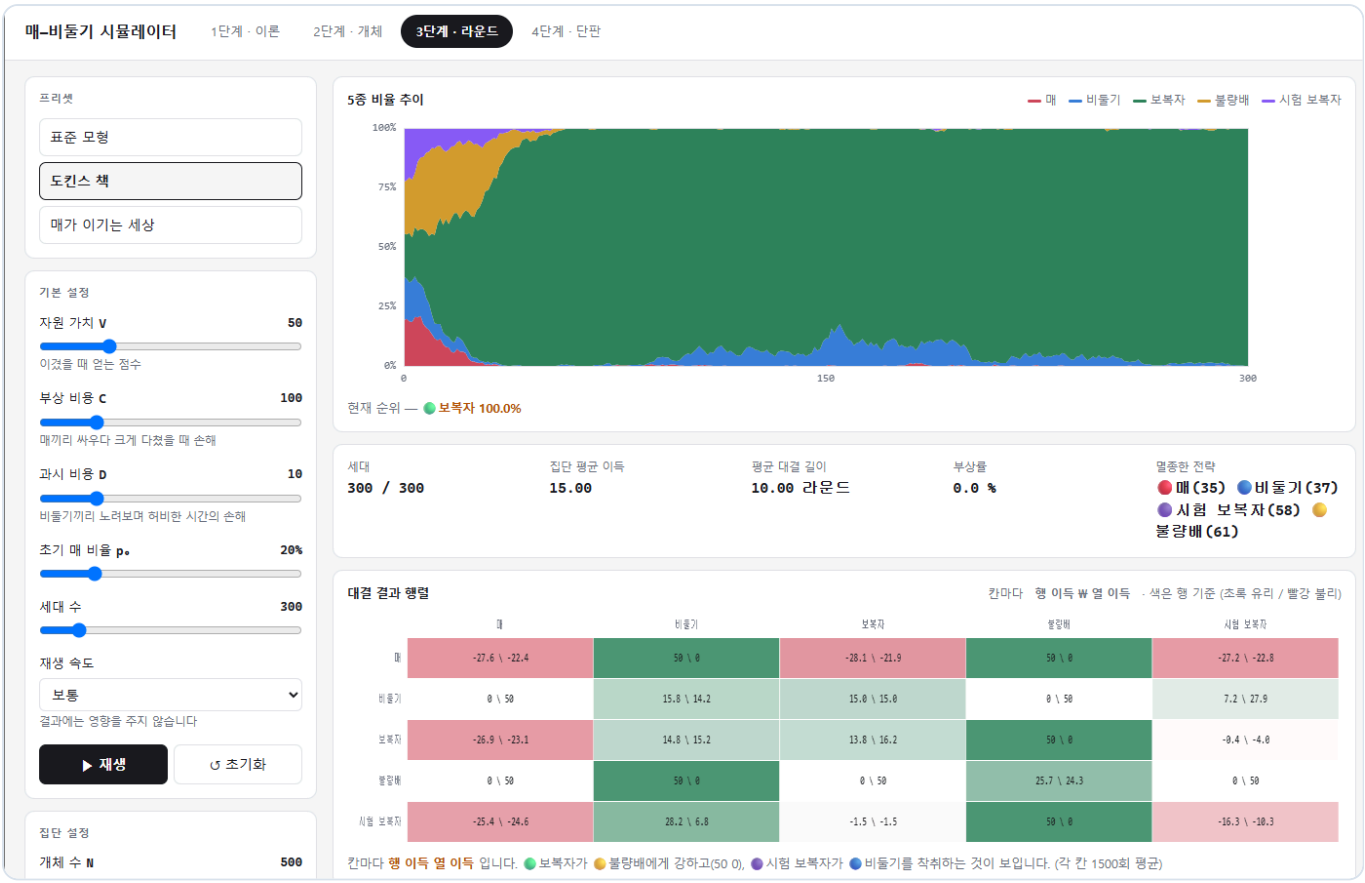
개체들이 정말 서로 맞붙는지 확인했고, 개체 수가 늘 5000마리에서 멈추는 이상한 결과를 보고 이유를 물었습니다. 코드 안에 개체 수 한계가 숨어 있었고, 저는 이 값을 화면에서 직접 바꿀 수 있게 했습니다.

> **내가 알게 된 것** — 결과가 이상하면 숨은 규칙이 있는지 물어봐야 한다. 내 시뮬레이터는 비율만 계산하는 것이 아니라 개체가 실제로 맞붙는다.

6 그렇게 내가 만든 시뮬레이터



2단계: 개체 500마리가 세대마다 실제로 맞붙고 번식합니다. 빨간 선(개체 시뮬레이션)이 회색 선(책의 계산)을 따라가며 58% 근처에 머뭅니다. 아래 칸 하나하나가 개체입니다.



3단계: 매파, 비둘기파, 보복자, 불량배, 시험보복자 다섯 전략의 경쟁. 세대가 지나며 보복자(초록)가 집단을 차지합니다. 책의 결론과 같습니다.

정리

- AI의 틀린 점을 거르려고 제가 계속 묻고 확인하다 보니, 한 번 더 공부하게 되었고 잘못 알던 것을 바로잡았습니다.
- 찾아봐도 없던 세대별·개체별 상황을 직접 만들어서, 궁금했던 것을 훨씬 구체적으로 이해했습니다.
- 이 방법은 『이기적 유전자』가 아니어도, 책이나 교과서에서 "정말 그럴까?" 궁금한 내용이면 누구나 쓸 수 있습니다.

참고: 이론 내용은 리처드 도킨스 『이기적 유전자』 5장을 바탕으로 했습니다. 사용한 AI: Claude Code (Anthropic). 화면은 직접 실행한 시뮬레이터를 캡처한 것이며 개인정보는 들어 있지 않습니다.