

NRF-간행물심의번호

NRF-20141223-1-20

NRF ISSN 2586-1131
ISSUE REPORT

2021_18호

병렬컴퓨팅 발전에 따른 계산과학 분야들의 발전과 확장, 그리고 연구 패러다임 변화

- I. 서론
- II. 계산과학 분야의 탄생과 발전, 그리고 연구 영역의 확장
- III. 연구 패러다임의 변화: GPU, AI, 데이터 기반 계산과학의
보편화
- IV. 국가별 사례
- V. 맺음말
- VI. 부록

CONTENTS

Ⅰ	서론	1
Ⅱ	계산과학 분야의 탄생과 발전, 그리고 연구 영역의 확장	4
	1. 계산과학 분야의 탄생	4
	2. 계산과학의 이론적 발전	7
	3. 병렬컴퓨팅 발전에 따른 연구 영역의 확장	14
Ⅲ	연구 패러다임의 변화: GPU, AI, 데이터 기반 계산과학의 보편화	21
Ⅳ	국가별 사례	24
	1. 미국	28
	2. EU	31
	3. 중국	32
	4. 일본	34
	5. 대한민국	36
Ⅴ	맺음말	43
Ⅵ	부록	46
	1. 계산과학의 정의	46
	2. 계산과학 연구 분야	48
	3. 계산과학 방법 및 알고리즘	50
	4. 계산과학 응용 분야	53
■	참고문헌	59

표 및 그림 목차

〈표 1〉 Editor들이 선정한 20세기 Top 10 알고리즘(배열 순서는 순위와 무관)	10
〈표 2〉 초고성능컴퓨터(Supercomputer) 개요	16
〈표 3〉 국가센터 초고성능컴퓨터 구축·운영 현황	18
〈표 4〉 HPL-AI Benchmark(2021년 6월)	26
〈표 5〉 주요국의 초고성능컴퓨팅 중점 활용 분야(미국)	29
〈표 6〉 주요국의 초고성능컴퓨팅 중점 활용 분야(EU)	31
〈표 7〉 중국의 슈퍼컴퓨터 쑤웨이 타이후 라이트의 대표 응용 사례들	32
〈표 8〉 인공지능(AI) 신기술 개발을 위한 액션 플랜(action plan)	35
〈표 9〉 주요국의 초고성능컴퓨팅 중점 활용 분야(일본)	36

[그림 1] 구글 딥마인드 알파폴드 2의 단백질 구조 예측 성능 수준 - 1	12
[그림 2] 구글 딥마인드 알파폴드 2의 단백질 구조 예측 성능 수준 - 2	12
[그림 3] 인간이 설계한 마이크로칩 평면도와 구글 AI가 생산한 평면도의 차이 예	13
[그림 4] 생체 분자 모델링 및 시뮬레이션 분야의 기대 곡선	17
[그림 5] 인공지능의 겨울(winter)과 봄(boom)	19
[그림 6] 전 세계 상위 200위 이내의 연산성능을 지니는 슈퍼컴퓨터들의 국가별 보유 현황 (2021년 6월 기준)	20
[그림 7] NVIDIA 1 GPU 노드(node)가 14~54 CPU 노드까지 대체하는 사례	23
[그림 8] 미·중·일·EU의 초고성능컴퓨팅 경쟁력 향상을 위한 정책역량 집중 사례	24
[그림 9] 전체 자원량 및 1위 시스템 자원량, 미·중·일·EU 및 한국의 총 자원량	25
[그림 10] 초고성능컴퓨팅 데이터 허브 체계	27
[그림 11] 연도별 국가센터 과제 지원율, 국내 자원 수요 및 실제 자원량	27
[그림 12] 국가별 1개 연구당 자원 배분 시간(2019)	28
[그림 13] 초고성능컴퓨팅 육성 로드맵 및 추진 전략 도식화(2021~2030)	37
[그림 14] 10대 전략분야별 유망기술 예시	38
[그림 15] 초고성능컴퓨팅 기반의 新서비스 예시	39
[그림 16] “산업수학”의 선정과제와 미선정과제의 키워드 클라우드	41
[그림 17] 산업수학 선정 과제 영문 키워드 네트워크 분석 클라우드	42
[그림 18] 산업수학 선정 과제 국문 키워드 토픽모델링	42

I. 서론

계산과학(computational science)은 고급 컴퓨팅 기술, 데이터 분석기술 등을 사용하여 복잡한 문제를 이해하고 해결해나가며 빠르게 성장하고 있는 다 학제 간 분야(multi- and interdisciplinary field)이다. 현재의 계산과학 수준은 과거에는 쉽지 않았던 여러 실험 결과들의 재현(reproduction), 다양한 현상들의 예측(prediction), 그리고 새로운 이론들의 보완(supplementation)을 훌륭히 수행해 내고 있는 수준이다. 하드웨어(hardware)의 발전으로 인해 과거에는 슈퍼컴퓨터(supercomputer)로만 가능했던 수많은 계산들이 오늘날에는 일반 개인용 컴퓨터(personal computer, PC)로도 계산해낼 수 있으며, 여러 계산과학적 시뮬레이션(simulation)들을 통해 산업 현장에서도 설계, 제품 개발 등의 영역들에서 비용 절감 및 시간 단축에도 큰 기여들을 하고 있다.

오늘날 초고성능컴퓨팅(supercomputing)은 게임의 룰(rule)을 바꾸는 증명된 기술로 받아들여지고 있다(미국경쟁력위원회, 2015). 계산과학 분야는 병렬컴퓨팅(parallel computing) 하드웨어와 관련 기술들의 발전과 함께 놀라운 속도로 성장해왔으며, 과학 분야를 구성하는 하나의 큰 축으로 자리매김하였다. 그뿐만 아니라, 인공지능(artificial intelligence, AI) 기술 활용의 보편화에 힘입은 현대 계산과학적 연구 트렌드는 거의 모든 연구 분야에 걸쳐 등장하고 있다. 오늘날 계산과학은 거의 모든 분야에 있어 하나의 기초 학문처럼 자리 잡고 있을 정도로 그 중요성이 두드러지고 있으며, 현재 빅데이터/인공지능 경제 가속화에 따른 계산 기술들의 선점 경쟁 격화 및 전략 무기화 움직임들이 미국, 중국, 일본, EU 등에 의해 활발히 일어나고 있다. 현대 계산과학적 방법들의 활용에 대한 자유도를 크게 높여주는 슈퍼컴퓨팅 인프라는 경제, 사회 전반의 혁신 플랫폼으로서 슈퍼컴퓨팅 인프라를 잘 갖추는 것은 국가 경쟁력을 높이기 위해 필수적이라 여겨지고 있다.

인공지능 기술들의 괄목할만한 성장을 필두로 계산과학 분야에서 불고 있는 새로운 바람과 함께, 새롭고 다양한 방법론들과 모델들이 전례 없는 수준의 빠른 속도로 개발되어 다양한 연구 분야 속에서 활발히 응용되고 있다. 하드웨어 기술의 발전으로 동일한 예산으로 과거보다 훨씬 더 좋은 성능의 고성능 컴퓨팅(high performance computing, HPC)이 가능해졌으며, 탐지

기술(detection techniques), 온라인 센서 네트워크(on-line sensor networks), 고해상도 이미징 기술(high-resolution imaging techniques) 등의 기술 발전은 더욱 풍부한 데이터를 기반으로 한 모델링 및 시뮬레이션을 가능하도록 하였다. 그리고, 병렬컴퓨팅 기반의 고성능컴퓨팅의 중요성이 시간이 지날수록 더욱 부각되고 있음에 따라 관련된 전자공학 및 반도체 제조기술(semiconductor manufacturing technology), 현대적 소프트웨어 기술들(modern software technologies)을 포함한 다양한 최첨단 기술들의 발전에도 영향을 미치고 있으며, 관련 최첨단 기술들의 발전은 서버(server), 스마트폰(smartphone), 네트워크 장비, 사물인터넷(internet of things, IoT) 등 4차 산업혁명(fourth industrial revolution, 4IR)과 관련된 산업 분야들의 기술 발전에까지 긍정적 영향을 주는 동시에 최첨단 기술 기반의 신산업 창출에 대한 파급효과까지 지닌다.

시간이 흐르면서 병렬컴퓨팅 및 계산과학과 관련된 하드웨어 기술 발전과 현대적인 계산 방법론들의 발전에 힘입어, 점차 과거에는 다룰 수 없었던 범위의 문제들을 다룰 수 있게 되었다. 그리고 풀리지 않았던 여러 분야들에 편재되어 있던 문제들 또한 해결할 수 있게 되었다. 그 문제들이 속하는 분야들은 과학, 공학, 의학 분야에서부터 사회학, 인문학 분야에 이르기까지 그 폭이 상당히 광범위하다. 위와 같은 학문 분야들에 속한 문제들 외에도, 시스템 하드웨어/소프트웨어 및 네트워크/데이터 관리 구성요소들을 개발하고 최적화하는 컴퓨터과학(computer science) 문제들 또한 계산과학 분야에 속하는 문제들이다. 최근 개정된 국가과학기술표준분류에 수학분야에 인공지능수학/데이터사이언스수학 소분류가 신설된 점도 이러한 빅데이터시대 인공지능기반 계산과학의 중요성이 반영된 것으로 이해된다. (이상남 외, 2020)

계산과학은 거의 모든 과학, 공학을 포함한 연구 분야들에 걸쳐 다양한 형태로 활용되고 있고, 정교한 수치 방법, 계산, 데이터, 네트워크 및 새로운 하드웨어들을 통합적으로 활용, 전반적인 연구 분야들에 걸쳐 복잡한 다중 규모, 다중 영역 문제들을 해결하는 혁신적인 첨단 학제 간 연구들이 수행되고 있다. 현재 다양한 분야에서 새롭게 개발되고 있는 모델들과 방법론들은 여러 학제 간의 지식 교환을 통해 시너지(synergy)를 발휘하며 새로운 형태의 연구 결과들의 창출을 촉진하고 있다. 더 나아가, 가까운 미래에는 현재 가장 빠른 일반 PC 5백만 대 이상의 성능을 지니는 엑사스케일(exascale)급 초고성능컴퓨팅 시스템(1초당 10^{18} 번 연산)이 본격적으로 활용되기 시작할 것으로 예상되며, 현재 미국, EU, 중국, 일본 등은 기존의 페타스케일(petascale)급에 비해 연산성능 및 집적도가 100배 수준으로 더 높고 전력효율 또한 3배 수준

더 높은, 엑사스케일급 슈퍼컴퓨팅 시스템들을 최대한 확충하기 위해 노력 중이다. 올해 6월 기준 HPL-AI 벤치마크 기준, 엑사스케일 급의 컴퓨팅 성능을 달성했던 슈퍼컴퓨터들은 일본의 후가쿠(Fugaku: 2.0 Eflop/s)와 미국의 서밋(Summit: 1.15 Eflop/s) 뿐이며, 가까운 미래로 갈수록 전 세계 각국은 엑사스케일급의 시스템을 갖추기 위한 경쟁을 가속화할 것으로 전망되고 있다. 관련하여, 2019년 기준 278억 달러 규모였던 고성능컴퓨터(high-performance computer)시장은 향후 2024년에는 429억 달러 수준으로 약 54.3% 정도 성장(연평균 약 9.0%) 할 것으로 전망되고 있다.

이러한 시대적 흐름 속에서, 대한민국은 연구 경쟁력의 신장을 위해 계산과학 기술 활용의 보편화를 극대화하기 위한 교육인프라 구축과 더불어 다른 선진국들에 크게 뒤처지지 않는 수준의 슈퍼컴퓨팅 인프라를 갖추어 계산자원 활용에 대한 접근성을 최대한 높일 수 있도록 하는 노력이 요구되고 있다. 대한민국 정부는 그간 초고성능컴퓨터법 제정(2011) 및 기본계획 수립(2013, 2018) 등을 통해 국가 초고성능컴퓨팅 육성체계를 구축하기 위해 노력해오고 있기는 하지만, 중장기적 실행전략 부족, 투자 정체 등의 한계로 자원 경쟁력이 최근 해마다 하락하고 있는 상황이며(슈퍼컴퓨터 TOP 500 국가별 순위 2015년 8위 → 2018년 이후 10위권 밖), 자체 기술력 확보와 활용성과 창출 또한 더딘 상황으로 올해 5월 말 비상경제 중앙대책본부(21-36)에 의해 진단된 상황이다.

이 리포트에서는 계산과학 분야의 탄생과 전반적인 발전 및 병렬컴퓨팅 기술 발전으로 인한 계산과학 분야의 확장 등에 대해 먼저 다루어 본 후, 과거 여러 학문 분야들이 어떤 흐름으로 계산과학을 접목, 새로운 계산과학 분야들을 만들어내기 시작하여 최근 GPU, 빅데이터, AI 기술들의 급격한 발전과 함께 연구 방식의 큰 패러다임(paradigm)의 변화까지 맞이하게 되었는가까지 살펴보고자 할 것이다. 관련하여, 미국, EU, 중국, 일본, 대한민국의 사례들에 대해서도 정리해 볼 것이다. 맺음말 이후에 포함될 부록에서는 계산과학의 일반론적인 부분들 - 정의, 연구 분야, 대표적 계산법 및 알고리즘, 응용(application) - 에 대해 전반적으로 다루어 둘 것이다.

II. 계산과학 분야의 탄생과 발전, 그리고 연구 영역의 확장

1 계산과학 분야의 탄생

인류에 의해 개발된 컴퓨터들은 방대한 양의 계산을 빠른 속도로 처리하기 위해 개발되었다. 그리고 그 방대한 양의 계산들이 필요했던 주된 이유는 과학, 수학 분야에서의 문제 해결을 위해서였다. 오늘날의 병렬컴퓨팅(parallel computing) 기술 기반의 초고성능컴퓨팅(supercomputing) 역시 수학, 과학, 공학 등 여러 분야에서의 여러 문제들과 난제들을 해결하기 위해 활용되고 있다. 미국의 물리학자 존 빈센트 아타나소프(John Vincent Atanasoff)에 의해 개발된 컴퓨터 역시 과학 컴퓨팅(scientific computing, SC)을 할 수 있기 위해서였고, 앨런 튜링(Alan Turing)에 의해 개발된 컴퓨터 또한 암호 해독 및 수학적 난제 해결(튜링의 리만 가설(Riemann hypothesis, RH) 증명 시도 등)을 위한 방대한 양의 계산들을 수행하기 위해서였다. 특히, 앨런 튜링과 그의 스승 존 폰 노이만(John von Neumann)에 의해, 현대 컴퓨터의 초기 버전이 탄생하게 되면서 이론 및 실험과학과는 다른, 계산과학이라는 새로운 분야가 탄생하여 자리 잡을 수 있게 되었다.

계산을 위한 기계는 기원전(Before Christ, B. C.)에 이미 존재했었다. 기원전 약 2700여 년 전 메소포타미아(Mesopotamia) 수메르(Sumer)에서 계산을 돕기 위해 주판(abacus, counting frames)이 사용되기 시작하였다고 추정되고 있다. 이후, 그리스(Greece)와 로마(Roma)를 거쳐 명나라 이후에 중국으로 전래, 개량된 것으로 알려져 있으며, 일본에 보급될 때 한 번 더 계량이 이루어졌다. 1970년대 전자계산기가 보급되기 전까지만 해도 주판은 계산을 위한 필수적인 사무용품으로 자리 잡고 있었다. 주판은 덧셈과 뺄셈뿐만 아니라 곱셈과 나눗셈도 처리할 수 있다.

1623년, 독일의 과학자 빌헬름 지카르트(Wilhelm Schickard)는 절친한 천문학자 요하네스 케플러(Johannes Kepler)가 6자리 덧셈과 뺄셈을 기계적으로 수행할 수 있는 기계를 개발, 이후 뉴턴의 이론 연구에까지 적지 않은 영향을 미치게 되는 케플러 법칙을 얻어내는 데 크게 이바지하였다. 1642년 블레즈 파스칼(Blaise Pascal)은 그의 아버지가 세울과 세금을 계산하는 것을 돕기 위한 것을 계기로, 상자 모양의 계산기를 발명하였다. 1725년 프랑스에서는 천공된 종이테이프(펀치카드)가 베틀(loom)에 붙여지는 형태로 직물 패턴을 짜는 일들을 보다 용이토록

하는 작업들을 시도하기 시작, 펀치카드의 교체를 통해 얻어지는 직물 패턴이 달라지도록 하는 시도들을 하였다. 초기에는 펀치카드 교체 시 직기의 디자인 변경도 필요했으나, 1801년 조셉 마리 자카드 (Joseph Marie Jacquard)는 직기의 기계적 디자인을 변경할 필요가 없는 자카드 직기(Jacquard loom)를 개발하여, 펀치카드만 디자인해서 교체해 주면 원하는 직조 패턴을 얻어낼 수 있도록 하였다. 이것은 프로그래밍 가능한 기계 제작의 출발점과 같은 의미를 지녔다. 1823년 영국의 수학자 찰스 배비지(Charles Babbage)는 증기 기관으로 구동되는 차분기관(differential engine)과 분석 기계를 구축하는 것을 시도하였는데, 차분기관을 통해 숫자의 차이를 비교하는 것을 시작으로, 기계로 하여금 펀치카드로 데이터를 입력, 저장, 분석할 수 있도록 그의 일생을 바치기도 하였다. 배비지의 도전이 성공하였다면 당시 프로그래밍 가능한 최초의 컴퓨터가 탄생하였을 것이고, 당시 그가 얻어내길 원했던 고정밀 로그 테이블 또한 인쇄해낼 수 있었을 것이다. 1890년 미국의 통계학자인 허먼 홀러리(Herman Hollery)는 펀칭을 사용하는 분류기를 개발하여 카드는 데이터를 저장, 기계를 카드를 감지하여 통계 데이터를 자동으로 기록할 수 있도록 하여 미국의 인구 조사 데이터를 처리하고 분석하는 데 필요한 시간을 약 8년에서 3년 이내로 단축할 수 있게 하였다. 1913년 MIT 교수 버니바 부시(Vannevar Bush)는 모터로 구동되는 최초의 아날로그 차동 분석기(differential analyser)를 발명하였다.

1936년 앨런 튜링(Alan Turing)은 컴퓨터 이론에 대한 기초 논문인 “On Computable Numbers, with Application to the Entscheidungs problem”을 발표하면서 현대 컴퓨터 이론에 필요한 중요한 개념들을 명확히 하였다. 튜링 기계(Turing machine)는 컴퓨터의 실현 가능성에 대한 것으로, 명제의 증명 가능성에 대한 연구에서 처음으로 이용된 개념이었는데, 본질은 괴델의 불완전성 정리(Gödel's incompleteness theorems): “페아노 공리계를 포함하는 모든 무모순적 공리계는 참인 일부 명제를 증명할 수 없으며, 특히 스스로의 무모순성을 증명할 수 없다.”의 또 다른 표현이었다. 튜링은 “주어진 프로그램이 해결하고자 하는 문제를 해결하는지 말해 줄 수 있는 일반화된 알고리즘이 존재하는가?”(정지 문제(halting problem))에 대한 질문에 일반화된 알고리즘은 존재하지 않음을 증명, 이는 “기계적인 방식으로는 수학의 모든 사실들을 만들어 낼 수 없다.”라는 ‘좌절적인 말’과도 같았다. 튜링의 이 증명은 괴델에 의해 수학계에 내리쳤던 이러한 ‘좌절’을 다시금 확인시켜 준 것이다. 하지만, 그 ‘좌절’을 증명하는 내용 속에 인류에게 정보 혁명을 안겨줄, 매우 중요한 도구의 설계도가 들어있었다. 튜링은 튜링 기계에 대해 다음과 같이 설명하고 있다.

“무한한 저장 공간은 무한한 길이의 테이프로 나타나는데 이 테이프는 하나의 기호를 인쇄할 수 있는 크기의 정사각형들로 쪼개져 있다. 언제든지 기계 속에는 하나의 기호가 저장되어 있고 이를 ‘읽은 기호’라고 한다. 이 기계는 읽은 기호를 바꿀 수 있는데, 그 기계의 행동은 오직 읽은 기호에 의해 결정된다. 테이프는 앞뒤로 움직일 수 있어서 모든 기호들은 적어도 한 번씩은 기계에게 읽힐 것이다.”

즉, 튜링 기계는 입출력 헤드의 상태와 헤드를 통해 테이프로부터 읽은 기호에 따라서 테이프에 입력된 기호가 바뀌거나 테이프가 앞뒤로 움직인다는 것으로, 튜링기계 하나는 하나의 알고리즘을 수행한다고 보고 이해하면 될 것이다. 더 나아가, 튜링은 하나의 기계로 모든 튜링 기계를 모사할 수 있는 보편 튜링 기계(universal Turing machine)라는 개념을 제안, 가상의 테이프에 알고리즘 자체를 저장하고 읽어내 사용할 수 있는 프로그램 내장형 컴퓨터(stored-program computer)의 개념을 제시하였다. 튜링이 프린스턴대학에서 박사 과정을 밟던 시기 그의 스승인 존 폰 노이만(John von Neumann) 교수와 친분을 다지게 되었었는데, 폰 노이만 교수는 당시 튜링 기계의 개념을 잘 알고 있었으며, 튜링이 완전히 실현하지 못했던 보편 튜링 기계를 결국에는 약간 변형된 형태로 구현해 내는 데 성공하게 된다. 이것이 오늘날 컴퓨터들의 첫 조상이며, 이후 컴퓨팅 하드웨어들의 발전은 오늘날에 이르기까지 놀라운 속도로 일어나게 된다.

또한, 앨런 튜링은 1950년 “Computing machinery and intelligence”라는 논문을 발표하며 인공지능(artificial intelligence, AI)에 대한 개념을 언급, 인공지능이라는 용어가 사용되기 전부터 이미 향후 인공지능이 광범위하게 사용될 것을 시대를 앞서 예측했던 것으로도 유명하다. 튜링은 튜링 테스트(Turing test)를 통해 지능이 인간의 본성이란 논쟁을 피하면서 지능에 관한 객관적이고 표준적인 시각을 제공하였고, 튜링 테스트는 현재에도 인공지능의 레벨을 판단하는 하나의 기준으로 활용되고 있다. 결국, 앨런 튜링은 인류에게 엄청난 혁명을 안겨다 준 컴퓨터와 인공지능의 아버지였던 셈이었고, 계산과학 분야의 탄생(birth of the field of computational science)과 혁명적 발전을 달성할 수 있게끔 과학 역사에 하나의 큰 획을 그었던 것이다. 앨런 튜링이 사망한 바로 다음 해인 1955년부터 인공지능(artificial intelligence)이라는 용어는 미국의 전산 및 인지 과학자 존 매카시(John McCarthy)의 논문에서부터 처음으로 사용되기 시작하였다.

1940년대 후반에 접어들며 디지털 전자 컴퓨터가 본격적으로 등장하기 시작하여 물리학 분야를 중심으로 활발히 활용되기 시작하였는데, 주로 선형 방정식(linear equation)들을

풀어내는 레벨로 제한적으로 활용되었던 한계가 있었다. 실제로 물리학에서 많은 중요한 방정식들은 비선형 방정식(nonlinear equation)들이었다는 것은 1800년대 후반부터 알려져 있었으며, 이들은 손으로 푸는 것이 불가능하였기에 컴퓨터가 이 문제를 해결해 줄 필요성이 대두되고 있었던 상황이었다. 1950년대부터 비선형 문제들의 해결에 도움을 주기 위한 여러 연구들이 있었으나 수많은 어려움들에 봉착하였다. 컴퓨터를 통해 근본적인 과학 문제들의 해결을 위해서는 비선형 측면에서 무언가 이루어져야 했었지만, 컴퓨터에게 이를 크게 기대할 수 없다는 분위기가 1970년대 중반까지 지배적으로 형성되었었다. 심지어 계산과학 분야 자체에 대해 부정적인 전망을 가졌었던 학자들도 다수 존재했었다. 1970년대 말에 접어들면서 컴퓨터 자원들을 사용하는 데 드는 비용이 크게 줄어들기 시작하여 이전보다 훨씬 더 저렴한 연구 비용으로, 더욱 다양한 도전적인 계산들을 시도해 볼 수 있게 됨에 따라 비선형 계산 문제들에 관한 연구들도 점차 활기를 되찾기 시작하였고, 기존의 이론 및 모델들에 기반한 흥미로운 계산과학적 연구 결과들도 점점 더 많이 발표되기 시작하였다. 마침내 1980년대에 접어들면서, 계산과학 분야는 본격적으로 기존의 이론과학 분야 및 실험과학 분야와 더불어, 하나의 독립적인 과학 분야의 한 축으로 인식되기 시작하였다.

2 계산과학의 이론적 발전

계산과학에 사용되고 있는 수학적 방법 및 이론들은 컴퓨터가 발명되기 훨씬 이전부터 얻어지기 시작하였고, 컴퓨터가 발명된 이후에는 컴퓨터로 계산된 수치들에서 얻은 통찰(insight)들과 함께 여러 계산과학적 방법과 이론들이 도출되었다. 우선, 컴퓨터가 발명되기 전 시기의 계산과학의 이론적 발전에 대해 살펴보도록 하면, 17세기 아이작 뉴턴(Issac Newton)에 의해 얻어진 뉴턴 방법(Newton's method), 몬테카를로 방법(Monte Carlo method)를 설계하는 데 사용할 수 있는 부푼의 바늘 문제(Buffon's needle problem)의 최초 제기, 18세기 토머스 심프슨(Thomas Simpson)에 의해 얻어진 심프슨 공식(Simpson's rule)과 레온하르트 오일러(Leonhard Euler)에 의해 얻어진 오일러 방법(Euler's method) 등에 대해 가장 먼저 떠올려볼 수 있다. 이들은 모두 오늘날 계산과학 영역에서도 기본이 되는 중요한 수치해석적 접근법들이다.

19세기에 접어들어서는 피에르시몽 드 라플라스 후작(Pierre-Simon, marquis de Laplace)에 의한 그램-슈미트 직교화(Gram-Schmidt orthogonalization)의 초기 공식이

얻어졌고, 유한 차분법(finite difference method, FDM)을 사용하는 계산기가 영국의 수학자 찰스 배비지(Charles Babbage)에 의해 고안되기도 하였으며, 인류 최초의 알고리즘으로 간주되는 예이다 러브레이스(Ada Lovelace)의 베르누이 수 생성 알고리즘(algorithm for generating Bernoulli numbers)도 이 시기에 등장하였다. 또한, 중요한 선형 다단계 방법들(linear multistep methods) 중 하나인 애덤스-배쉬포스 방법(Adams-Bashforth method)과 강한 대각지배행렬(strictly diagonally dominant matrix)로 이루어진 연립 일차방정식(system of linear equations)에서 반복법(iterative method)의 수렴성(convergence)을 보증하는 연립 일차방정식의 풀이법인 독일의 수학자 카를 구스타프 야코프 야코비(Carl Gustav Jacob Jacobi)의 야코비 방법(Jacobi method), 그리고 연립방정식에 대응하는 행렬을 두 개의 삼각 행렬(triangular matrix)로 분리한 뒤 해를 반복적으로 계산해 수렴된 해를 얻어 연립방정식을 수치적으로 풀어내는 가우스-자이델 방법(Gauss-Seidel method)도 19세기에 얻어졌다.

20세기 초기에 접어들면서, 미분 방정식(differential equation)의 적분(integration)을 근사하기 위한 룬게-쿠타 방법(Runge-Kutta method), 에르미트 행렬(Hermitian matrix)과 양의 정부호행렬(positive-definite matrix)의 분해에 활용되는 솔레스키 분해(Cholesky decomposition), 수치적 외삽에 활용되는 리처더슨 외삽(Richardson extrapolation)이 얻어졌고, 1920년대에는 기상 예보(weather forecast)를 위한 수치적 계산을 시도하기 시작하며 계산과학의 이론적인 발전들이 실제 현상들의 예측에 활용되기 시작하며 다양한 이론적 모델들과 계산과학적 방법들이 쏟아져 나오기 시작하였다. 1927년 더글라스 하트리(Douglas Hartree)는 훗날 하트리-포크 방법(Hartree-Fock method)으로 알려진 최초의 제일원리(First-principles) 양자화학 계산(quantum chemical calculation) 방법의 이론적 기반을 구축하였으며, 페르미(Fermi)의 로마 물리학 연구 그룹은 부폰의 바늘 문제(Buffon's needle problem)에서의 아이디어에 기반을 둔 통계적 알고리즘을 개발, 몬테카를로 방법(Monte Carlo method) 개발의 초석을 다졌다. 그리고 미국의 수학자 클로드 섀넌(Claude Shannon)은 전기 회로 사용을 통한 부울 대수 수행법을 소개하였다.

1940년대에는 인류 최초의 전자식 디지털 컴퓨터(electronic digital computer)인 아타나소프-베리 컴퓨터(Atanasoff-Berry computer)가 개발되었다. 1940년, 미국의 물리학자 존 빈센트 아타나소프(John Vincent Atanasoff)의 논문인 "Computing Machine for the Solution of Large Systems of Linear Algebraic Equations"에서는 손으로 푸는 데 최소

125시간이 걸릴 20개의 선형 방정식 시스템(a system of 20 linear equations)을 고안한 컴퓨터로 약 12시간 만에 풀어낼 수 있다고 추정했었다. (오늘날의 컴퓨터는 0.01초 이내에 10,000개 규모의 선형 방정식 시스템도 어렵지 않게 풀어낼 수 있다.) 그리고, 제2차 세계대전 중 앨런 튜링(Alan Turing)에 의해 이론적으로 고안되어 개발된 암호 해독용 컴퓨터는 당시 독일군의 에니그마(Enigma)로 인해 생성된 암호들을 해독, 연합군의 승리에 크게 기여하였다. 앨런 튜링은 현재의 컴퓨터가 탄생할 수 있도록 한, 보편 튜링 기계(universal Turing machine)의 개념을 이론적으로 제시하였다.

이후, 1947년 몬테카를로 시뮬레이션(Monte Carlo simulation) 방법이 존 폰 노이만(John von Neumann) 등에 의해 얻어졌으며, 선형계획 모델을 가우스 오퍼레이션을 통해 손쉽게 풀 수 있는 심플렉스법(simplex method)도 같은 해에 조지 댄치그(George Dantzig)에 의해 얻어졌다. 또한, LU-분해법(LU-Decomposition method)이 앨런 튜링(Alan Turing)에 의해 얻어지기도 하였으며, 필립스 유압 컴퓨터(Phillips hydraulic computer)로 알려졌고, 모니악 유압 컴퓨터(MONIAC hydraulic computer)가 개발되어 최초의 하이드로 시뮬레이션(hydro simulation)이 미국 로스앨러모스(Los Alamos)에서 수행되었다.

1950년대에 접어들며 앨런 튜링은 튜링 테스트(Turing test)의 개념과 함께 인공지능에 대한 개념을 소개하였다. 또한, 컴퓨터를 통한 날씨 예보를 성공적으로 수행할 수 있게 되기 시작하였으며, 크릴로프 부분 공간 반복법(Krylov subspace iteration methods)의 연구를 통한 란초스 방법(Lanczos method)이 얻어지고, 이후 예조건화(preconditioning)를 포함하는 데이비드슨-란초스 방법(Davidson-Lanczos method)도 얻어지게 되어 향후 행렬 고유치 문제들을 수치적으로 효과적으로 풀어내는 데 막대한 이바지를 하게 되었다. 그리고 하우스홀더 변환(Householder transformation)도 이 시기에 얻어졌다. 또한, 물리학자 베르니 앨더(Berni Alder)와 토마스 에버렛 웨인라이트(Thomas Everett Wainwright)에 의하여 분자동역학(molecular dynamics)이 고안되었으며, 캘리포니아 산호세 IBM 연구 센터에서 가장 오래된 프로그래밍 언어인 포트란 언어(FORTRAN language)와 컴파일러(compiler)도 1950년대에 개발되었다. 1955년에는 인공지능(artificial intelligence)이라는 용어가 존 매카시(John McCarthy)의 논문에서부터 처음으로 사용되기 시작하였고, 1957년 프랑크 로젠블라트(Frank Rosenblatt)는 현대의 인공 신경망을 구성하는 가장 기본 요소인 퍼셉트론(perceptron)의 개념을 고안하였다.

〈표 1〉 Editor들이 선정한 20세기 Top 10 알고리즘(배열 순서는 순위와 무관)

알고리즘 이름
Metropolis Algorithm for Monte Carlo
Simplex Method for Linear Programming
Krylov Subspace Iteration Methods
The Decompositional Approach to Matrix Computations
The Fortran Optimizing Compiler
QR Algorithm for Computing Eigenvalues
Quicksort Algorithm for Sorting
Fast Fourier Transform
Integer Relation Detection
Fast Multipole Method

출처: Society for Industrial and Applied Mathematics(SIAM) 뉴스(2000년 5월 16일)

1960년대에도 계산과학계에 커다란 이론적 진보들(theoretical advances)이 있었다. 1960년에는 유한요소해석법(finite element method)이라는 용어가 등장하며 활용되기 시작하였다. 그리고 1961년에는 영국의 과학자 존 프랜시스(John Francis)와 러시아 수학자 베라 쿠블라노브스카야(Vera Kublanovskaya)에 의해 QR 분해(QR factorization)법이 얻어졌고, 1963년에는 나비 효과(butterfly effect)와 혼돈 이론(chaos theory)의 이론적 개념이 등장하였다. 1964년, 1998년 노벨 화학상(Nobel Prize in Chemistry)을 수상토록 한, 밀도범함수 이론(density functional theory)의 기초가 되는 콘-삼 방정식(Kohn-Sham equation)이 이론 물리학자 월터 콘(Walter Kohn), 루 주 샴(Lu Jeu Sham), 피에르 오양베르그(Pierre Hohenberg)에 의해 얻어졌고, 같은 해 물리학자 아니수르 라흐만(Aneesur Rahman)에 의한 분자 동역학(molecular dynamics)도 이 시기에 독립적으로 얻어졌다. 1965년에는 제임스 윌리엄 쿨리(James William Cooley)와 존 튜키(John Tukey)에 의해 고속 푸리에 변환(fast Fourier transform)법 또한 얻어졌다. 경제학 문제들의 수치적 회귀(numerical regression) 분석도 1960년대부터 본격적으로 시작되었다.

1970년대 접어들면서 망델브로(Mandelbrot) 집합에서 발견되는 자기 유사성(self similarity)을 설명하기 위해 프랙털(fractal)이라는 개념이 만들어졌고, 프랙털은 계산과학적 방법들을 통해 광범위하게 시각화되며 탐구되기 시작하였다. 1973년에는 블랙-숄츠 모델(Black-Scholes model)이 알베르트 아인슈타인(Albert Einstein)의 브라운 운동(Brownian motion) 방정식으로부터 영감을 받아 얻어져 1997년 노벨 기념 경제학상(Nobel Memorial Prize in Economic Sciences)을 수상토록 하였다. 그리고 컴퓨터 계산을 통해 증명된 최초의 수학 정리인 4색 정리(four colour theorem)는 1977년 최초로 계산과학적 방법을 통해 증명되었다. 또한, 1980년대에는 대수적 계산(algebraic computation)을 수행해낼 수 있는 프로그램의 개발에도 큰 진전이 있었다. 이론 물리학자 스테판 울프람(Stephen Wolfram)은 컴퓨터를 통해 수치해석의 세부적인 요소들에 대해 신경 쓸 필요가 없는, 정확한 대수적 계산들을 해낼 수 있는 컴퓨터 대수(computer algebra) 소프트웨어 시스템인 SMP(Symbolic Manipulation Program)를 개발하여, 컴퓨터에 직접 수식을 넣어 기호 계산(symbolic computation)을 수행할 수 있도록 하였다. 스테판 울프람의 SMP는 현존하는 가장 강력한 기호 계산(symbolic computation) 소프트웨어인 맵스매티카(Mathematica)의 전신이다.

1980년대 이후에도 여러 계산과학적 모델들과 계산법들이 이론적으로 고안되어 활용되기 시작하였다. 가장 괄목할만한 부분은 컴퓨터 성능 발전에 힘입어 분자 생물학(molecular biology), 유전학(genetics) 분야 등에서의 컴퓨터를 활용한 단백질 구조 예측, 유전자 염기서열 분석, 인간 게놈 프로젝트(Human Genome Project) 등과 같은 보다 복잡도가 높은 계산과학적 문제들을 다룰 수 있게 된 부분, 그리고 데이터 과학(data science) 및 인공지능 이론들(theories of artificial intelligence)의 괄목할만한 성장 덕에 전산 생물학(computational biology) 분야를 포함한 계산과학 분야의 전반적인 큰 성장을 이루어낼 수 있게 된 부분이다. 최근 구글 딥마인드(Google DeepMind)의 AI인 알파폴드 2(AlphaFold 2)로 인간 단백질 2만 개 구조를 모두 예측, 36만 5,000개 이상 단백질의 3차원 구조를 정확히 예측하는 데 성공한 연구 결과가 2021년 7월 22일 국제학술지 네이처(Nature)에 발표된 바 있다. 구글의 알파폴드 1과 알파폴드 2는 아직 실험적으로 밝혀지지 않은 SARS-CoV-2 변이 바이러스들의 일부 단백질 구조들을 예측하는 데에도 활용되고 있다.

[그림 1] 구글 딥마인드 알파폴드 2의 단백질 구조 예측 성능 수준 - 1

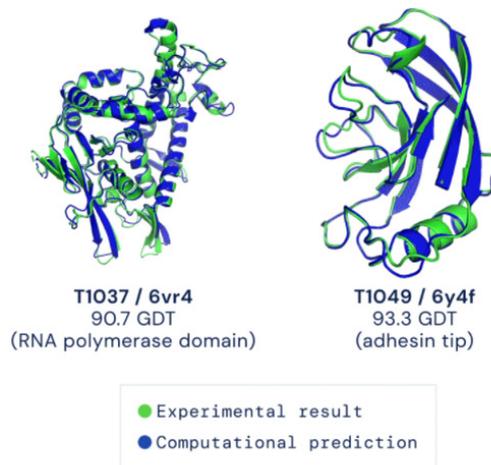
Median Free-Modelling Accuracy



Improvements in the median accuracy of predictions in the free modelling category for the best team in each CASP, measured as best-of-5 GDT.

출처: <https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>

[그림 2] 구글 딥마인드 알파폴드 2의 단백질 구조 예측 성능 수준 - 2

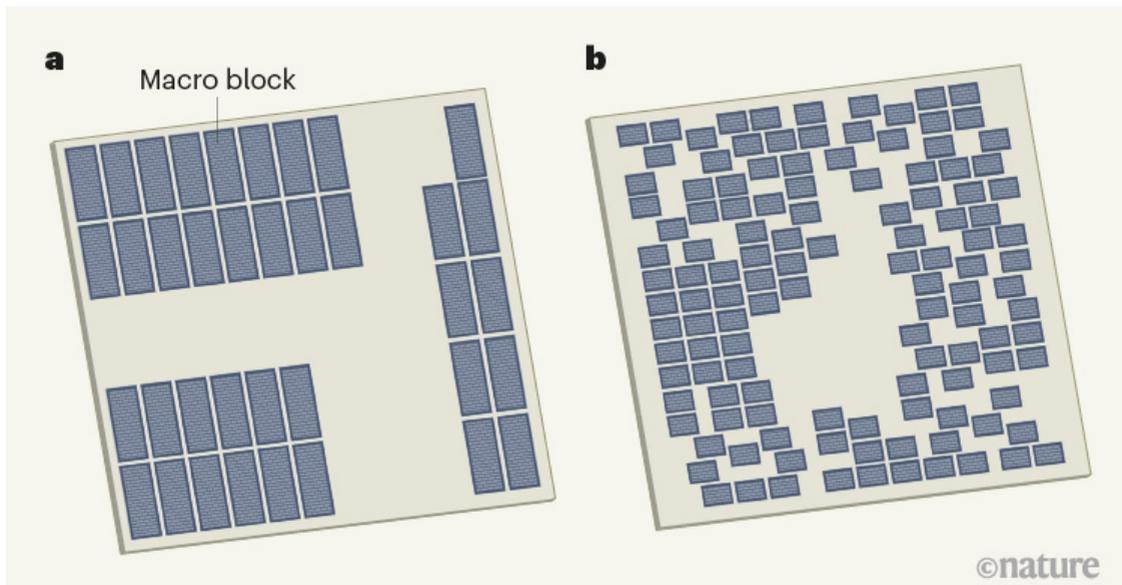


Two examples of protein targets in the free modelling category. AlphaFold predicts highly accurate structures measured against experimental result.

출처: <https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>

그뿐만 아니라, 구글의 AI는 반도체(Semiconductor) 설계 전문가가 수개월에 걸쳐 설계한 반도체보다 더 우수한 성능을 낼 수 있는 반도체를 단 몇 시간 만에 인간이 상상할 수 없었던 획기적인 방식으로 설계해 내었다는 연구 결과도 2021년 6월 국제학술지 네이처에 발표되기도 하였다.

[그림 3] 인간이 설계한 마이크로칩 평면도와 구글 AI가 생산한 평면도의 차이 예



출처: <https://www.nature.com/articles/d41586-021-01515-9>

이처럼, 다양한 계산법들과 수학적 모델들의 이론적 도출을 포함한 계산과학의 이론적 발전은 수학을 기반으로 하여 컴퓨터가 존재하기 훨씬 더 과거부터 현재까지 꾸준히 이루어져 왔다. 그리고 오늘날의 계산 과학자들은 인공지능의 활용을 통한 계산과학 분야 자체의 패러다임(paradigm)이 변하고 있음을 실감하고 있으며, 이에 따라 인공지능의 발전 및 효과적 활용을 위한 이론적 연구들에도 큰 노력을 기울이고 있는 상황이다.

대규모의 병렬컴퓨팅(parallel computing)을 수행하는 슈퍼컴퓨터 또는 초고성능컴퓨터(supercomputer)는 유용한 연구 통찰력(scientific insight)들을 제공해 줄 수 있는 좋은 도구로 진화해왔고, 병렬컴퓨팅 하드웨어 및 기술들이 발전해 옴에 따라 다룰 수 있는 문제들의 복잡도(complexity) 수준 또한 발맞춰 상승, 계산과학 분야의 연구 영역 또한 확장되었다. 특히, 그래픽 처리 장치(graphics processing unit, GPU) 병렬컴퓨팅의 놀라운 발전 및 데이터 급증과 더불어, 인공지능(artificial intelligence, AI)의 발전 및 고도화 등이 이에 힘입어 함께 이루어짐에 따라 지능정보사회 확산이 급속화되고 있으며, 이에 따라 고성능컴퓨터(high-performance computer, HPC) 및 초고성능컴퓨터는 핵심 인프라로 대두되고 있다.

병렬컴퓨팅의 발전을 논하기에 앞서, 오늘날 우리가 사용하는 컴퓨터의 시초와 간략한 발전 과정들에 대해 짚어보도록 하자. 오늘날 컴퓨터의 첫 번째 버전은 1930~40년대 앨런 튜링(Alan Turing)과 존 폰 노이만(John von Neumann)에 의해 고안되고(designed) 만들어졌다(created). 그리고 1950년대부터 컴퓨팅 하드웨어(computing hardware)의 괄목할만한 발전이 시작되었다. 컴퓨터의 역사는 ‘세대(generation)’를 통해 나누어진다.

제1세대 컴퓨터(first-generation computer)는 정보(information)를 저장하기 위한 소자로 진공관(vacuum tube, electron tube)을 사용하였고, 주기억장치(main memory)는 자기 드럼(magnetic drum)이 사용되었으며, 프로그램은 기계어(machine language)로 작성되었다. 제2세대 컴퓨터(second-generation computer)는 정보를 저장하기 위한 소자로 진공관보다 작고 안정적인데 빠르고 저렴한 트랜지스터(transistor)를 사용하였으며, 제3세대 컴퓨터(third-generation computer)는 집적회로(integrated circuit, IC)를 사용하여 정보를 저장하였다. 집적회로의 발명은 컴퓨터의 소형화와 가격 하락을 불러왔고, 마이크로프로세서(microprocessor)의 개발 계기가 되었다. 제3세대 컴퓨터는 운영체제(operating system, OS)를 통해 컴퓨터의 작업이 관리되었고, 다중 프로그래밍(multi-programming)과 시분할(time-sharing) 등을 가능케 하였다. 제4세대 컴퓨터(fourth-generation computer)는 오늘날의 컴퓨터이다. 제4세대 컴퓨터에는 하나의 칩에 수천 개에서 수백만 개 이상의 전자회로 소자를 집적시킨 고밀도 집적회로(large scale integration, LSI), 초고밀도 집적회로(very large scale integration, VLSI), 그리고 중앙처리장치(central processing unit, CPU)

마이크로프로세서(microprocessor)가 사용된다. LSI 기술 발전을 통한 마이크로프로세서의 개발은 개인용 컴퓨터(personal computer, PC)가 대량으로 보급될 수 있도록 하였다. 제4세대 컴퓨터의 보급과 관련하여, 스티브 잡스(Steve Jobs)의 애플(Apple)은 사용자들이 스스로 응용 프로그램들을 만들 수 있는 애플-2(Apple-2)를 1977년에 출시하여 큰 성공을 거두었고, IBM은 애플에 대응하기 위해 컴퓨터의 설계에 대한 정보의 공개와 함께 IBM PC를 1981년에 출시하여 PC 시장의 주도권을 잡게 되었다. 이때, IBM PC의 운영체제는 빌 게이츠(Bill Gates)의 마이크로소프트(Microsoft)가 개발한 운영체제인 MS-DOS가 사용되었다. 1984년 애플은 마우스와 그래픽을 이용하는 그래픽 사용자 인터페이스(graphic user interface, GUI) 기반 운영체제에 의해 구동되는 매킨토시(Macintosh) 컴퓨터를 출시하였는데, 매킨토시 컴퓨터는 오늘날 우리가 사용하고 있는 GUI 기반 운영체제들의 첫 조상 격이었다.

오늘날의 컴퓨터인 제4세대 컴퓨터의 성능은 CPU의 성능에 달려있었다. CPU의 성능 향상은 대체로 인텔의 고든 얼 무어(Gordon Earle Moore)가 제시한 무어의 법칙(Moore's law) - 컴퓨터의 성능은 일정 시기마다 배가하며 기하급수적으로 증가한다는 법칙 - 에 따라 이루어져 왔다. 계산과학 분야에서 사용해오던 컴퓨터들의 성능 또한 사용되는 CPU들의 성능에 따라 결정되었다. 그리고 큰 규모의 연산들을 필요로 하는 계산과학 문제들을 다룰 때는 다수 개의 CPU들(multiple CPUs)을 동시에 사용하여 계산을 수행할 수 있도록 하는 CPU 병렬 연산(parallel computing) 기술들이 개발되어 활용되었다. 일반적으로, 병렬컴퓨팅 시스템의 규모가 당대 최상급의 연산 처리 능력을 갖출 정도로 충분히 큰 컴퓨터를 우리는 슈퍼컴퓨터(supercomputer)라고 부른다. 슈퍼컴퓨터의 개념은 절대적 기준이 아닌 상대적 기준의 개념으로, 슈퍼컴퓨터와 같은 부류의 컴퓨터들을 단순히 고성능컴퓨터(High-Performance Computer, HPC)라 부르기도 하지만, 최근 5월 발에 발행된 국가초고성능컴퓨팅 혁신전략 문서에서 기술하고 있는 슈퍼컴퓨터, HPC 등의 정의는 다음 표의 내용과 같다.

〈표 2〉 초고성능컴퓨터(Supercomputer) 개요

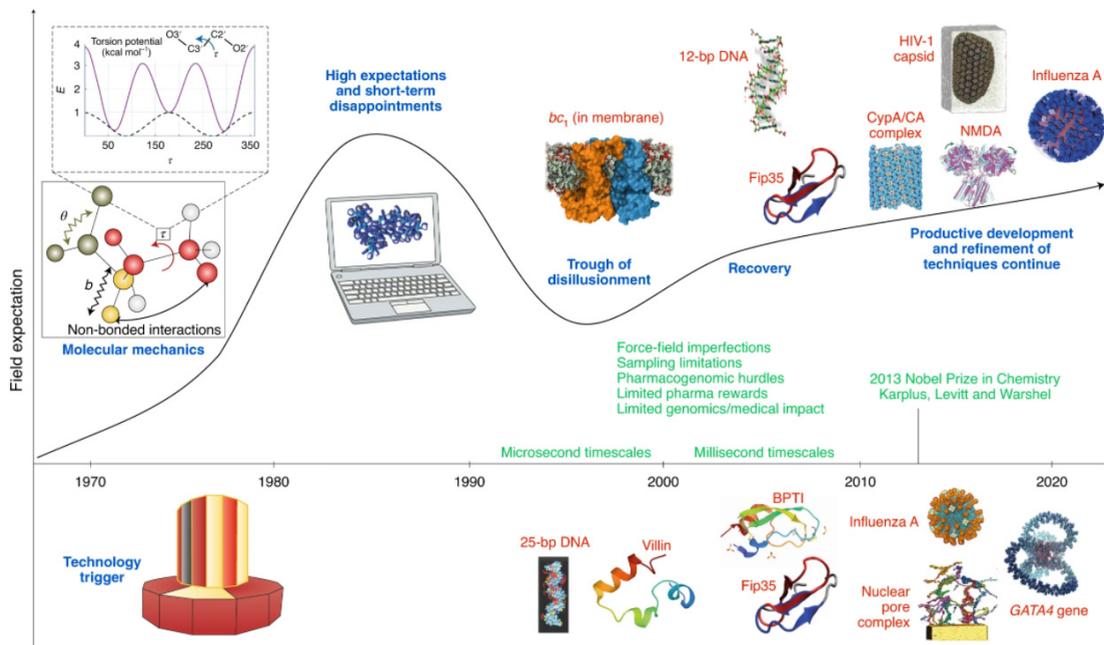
명 칭	세부 기준	비 고
탑 500 초고성능컴퓨터 (Top500 Supercomputer)	매년 두 차례(6월, 11월)에 걸쳐 발표하는 전 세계에서 성능이 가능 뛰어난 500대의 초고성능컴퓨터	
초고성능컴퓨터 (Supercomputer)	150만 달러를 초과하는 고성능컴퓨터	
고성능컴퓨터 (High Performance Computer, HPC)	일반컴퓨터보다 뛰어난 계산·메모리·데이터를 기반으로 과학·공학·데이터분석 등에 사용되는 서버급 이상 컴퓨터	

※ 서버 구분(Intersect360 Research) : ① Entry(<50K\$), ② Midange(50K\$~250K\$), ③ High-End(250K\$~1.5M\$), ④ Supercomputer(1.5M\$>)

출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

최초의 슈퍼컴퓨터는 1964년 미국의 CDC 6600으로 알려져 있다. 그리고 대중적으로 알려지기 시작한 슈퍼컴퓨터는 1985년 미국의 Cray-2이다. 대한민국 최초의 슈퍼컴퓨터는 1988년 11월에 도입된 2GFlops의 성능을 갖던 Cray-2S였으며, 이후 한국과학기술정보연구원(KISTI)과 기상청을 중심으로 슈퍼컴퓨터가 도입되어 활용되기 시작하였다. 근래에는 국내 다수의 대학 및 연구기관들도 HPC들을 도입하여 연구에 활용하고 있다. 슈퍼컴퓨터를 직접 도입하여 활용할 수 없는 경우에도 일정 비용 지불을 통해 KISTI 등의 슈퍼컴퓨터를 직접 사용하거나, 또는 구글(Google), 아마존(Amazon), 마이크로소프트(Microsoft) 등에 의해 제공이 되고 있는 클라우드 계산 플랫폼(cloud computing platform)들을 활용하여 거대 계산과학 연구 수행이 가능하다. 가용 계산 자원이 증가할수록, 아래 그림과 표에서의 예들과 같이 더욱 거대한 규모의 시뮬레이션이 가능해지며, 이는 다른 계산과학 분야의 시뮬레이션들에 대해서도 동일하게 적용된다.

[그림 4] 생체 분자 모델링 및 시뮬레이션 분야의 기대 곡선



* 이 분야를 주도하는 데 도움이 된 기술 혁신에는 분산 계산과 생체 분자 계산을 위한 GPU 사용의 출현이 포함되었다. (원문: Technological innovations that have helped drive the field include distributed computations and the advent of the use of GPUs for biomolecular computations.)

* 위 그림에 묘사된 랜드마크 시뮬레이션들의 규모(해가 지남에 따라 큰 증가 추세)

25-bp DNA (5 ns, ~21,000 atoms)

villin protein (1 μ s, 12,000 atoms)

bc1 membrane complex (1 ns, ~91,000 atoms)

12-bp DNA (1.2 μ s, ~16,000 atoms)

Fip35 protein (10 μ s, ~30,000 atoms)

Fip35, bovine pancreatic trypsin inhibitory (BPTI) proteins (100 μ s for Fip35, 1 ms for BPTI, ~13,000 atoms)

nuclear pore complex (1 μ s, 15.5 million atoms)

influenza A virus (1 μ s, >1 million atoms)

N-methyl-D-aspartate (NMDA) receptor in membrane (60 μ s, ~507,000 atoms)

tubular cyclophilin A/capsid protein (CypA/CA) complexes (100 ns, 25.6 million atoms)

HIV-1 fully solvated empty capsid (1 μ s, 64 million atoms)

GATA4 gene (1 ns, 1B atoms)

influenza A virus H1N1 (121 ns, ~160 million atoms)

출처: <https://www.nature.com/articles/s43588-021-00060-9>

〈표 3〉 국가센터 초고성능컴퓨터 구축·운영 현황

구분	운영기간*	성능 (TFlops)	‘초고성능 컴퓨터 탑 500’ 순위**	구축예산 (억원)	주요 활용 성과
1호기	’88~’93(5년)	0.002	-	180	<ul style="list-style-type: none"> 국산 자동차 설계 지원으로 개발기간·비용 단축 과학적 기상예보 시작
2호기	’93~’01(8년)	0.015	23위	260	<ul style="list-style-type: none"> 액체로켓 엔진 연소실 내부 시뮬레이션으로 액체로켓 효율 향상 탄소나노튜브를 이용한 평면 디스플레이 제작
3호기	’01~’09(8년)	4.6	42위	594	<ul style="list-style-type: none"> 세계 최대 우주 시뮬레이션을 통해 우주의 생성·진화 과정 규명
4호기	’08~’18(10년)	363.6	14위	732	<ul style="list-style-type: none"> 대체작물 미래 생산성 예측을 통한 기후변화 대응 식량생산 정책 수립 지원
5호기	’18~현재	25,700	11위	908	<ul style="list-style-type: none"> 초거대규모 난류 열유동 해석을 통해 난류에너지 생성·변화 과정 규명

* 해외 주요국 시스템별 평균 운영 기간: 4~5년

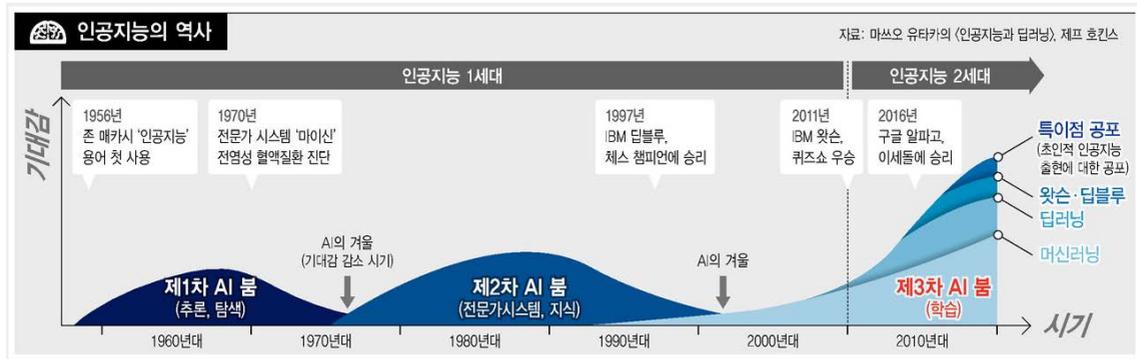
** 구축 완료 당시 기준

출처: 4차 산업혁명 권텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

병렬처리 기술들에 대한 중요성은 2005년도에 접어들면서부터 더욱 부각되기 시작하였다. 단일 CPU 코어의 성능이 4GHz 대를 넘어서기 어려워진 한계에 봉착했기 때문이다. 이후, 듀얼코어(dual-core), 쿼드코어(quad-core)와 같은 개념들이 생겨나면서 멀티코어(multi-core) CPU들이 본격적으로 개발, 보급되기 시작하였고, 이에 따라 새롭게 발전하고 있는 CPU들을 사용하여 연산 속도 향상을 이뤄내기 위해서는 병렬처리 기술들의 활용이 불가피하게 되었다.

또한, 2010년대에 접어들어서는 CPU의 활용만으로는 HPC 성능 발전에 한계가 있다는 인식이 커지게 되었으며, 2010년대 초중반부터 멀티 GPU들을 활용한 HPC들이 구성되어 CPU만으로 구성된 HPC 성능의 한계를 극복하기 위한 노력이 본격적으로 시작되었다. 당시, 멀티 GPU들의 활용을 통한 병렬 연산은 같은 비용으로 멀티 CPU들만으로 병렬 연산을 수행하였을 경우보다 주어진 계산과학 문제의 성격에 따라 거의 10배 내외의 연산 속도 향상을 보여주기도 하였다. 위 [그림 4]에서 언급된 전산 생물학 모델링 및 시뮬레이션의 발전에도 GPU의 발전이 크게 이바지하였고, 2010년도 초까지 지속되어 온 AI 겨울(AI winter) 또한 컴퓨팅 연산성능의 비약적 발달과 함께, 이 시기에 접어들어 비로소 극복되었다.

[그림 5] 인공지능의 겨울(winter)과 봄(boom)



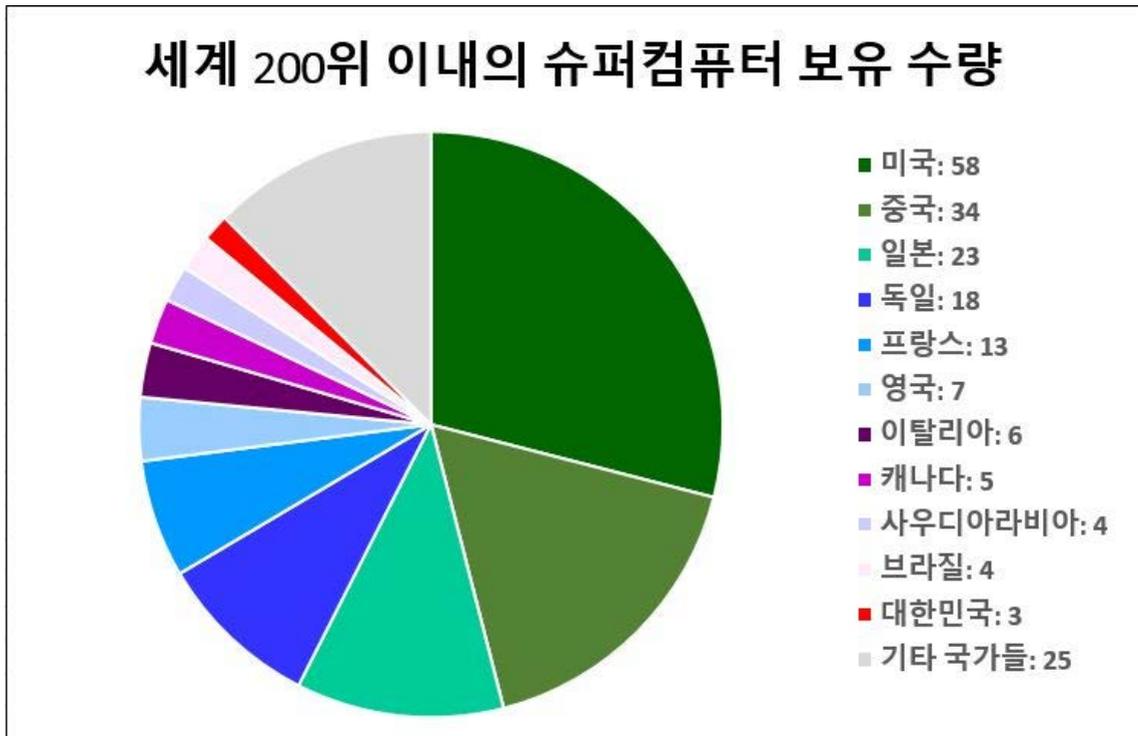
출처: '고양이 얼굴' 알아본 인공지능, 움직임까지 예측(2017), 한겨레 신문

2016년 전 세계에 이름을 알렸던 구글(Google)의 알파고(AlphaGo) 역시 멀티 GPU들의 병렬 연산성능의 극대화를 통해 당대 최고의 인공지능 성능을 구현해낼 수 있었다. 현재 전 세계 상위 500대의 슈퍼컴퓨터들 중 상당수가 엔비디아(NVIDIA)의 GPU들을 활용하여 연산성능을 극대화하고 있다. 또한, 연산성능을 더욱 효율적으로 향상시키기 위해, 구글은 텐서 처리 장치(tensor processing unit, TPU)를, 퀄컴(Qualcomm)은 신경망 처리 장치(neural processing unit, NPU)를 개발하고 있다. 하지만, 예외적으로 2021년 6월 기준에서의 세계 1위의 연산성능을 갖는 슈퍼컴퓨터는 64비트 ARM 아키텍처 기반의 마이크로프로세서인 A64FX 48C 2.2GHz들로 구성된 일본 리켄(RIKEN)의 슈퍼컴퓨터 후카쿠(Fugaku)이다. ARM 아키텍처는 저전력, 고효율에 초점을 맞추고 있는 특징이 있으며, 특히 A64FX는 512비트 벡터 구현이 포함된 ARMv8.2-A Scalable Vector Extension SIMD(single instruction multiple data) 명령어 세트를 사용하는 최초의 프로세서라는 특징이 있다. 후카쿠의 이론적 최대 연산성능은 537,212 TFlop/s이다.

한편, 대한민국은 2021년 6월 기준 상위 100위 안에 드는 슈퍼컴퓨터들을 일본보다 15대 적은(일본은 상위 100위 안에 총 18대 보유 중) 총 3대를 보유하고 있는 상황이며, 그 중 가장 빠른 슈퍼컴퓨터는 23위를 기록하고 있는 Maru와 Guru이다. 이들의 이론적 최대 연산성능은 일본 후카쿠의 약 1/20 수준인 25,495.1 TFlop/s이다. 현재 대한민국 최고의 슈퍼컴퓨터인 Maru와 Guru의 연산 속도는 대한민국 최초의 슈퍼컴퓨터인 Clay-2S 연산 속도보다 약 1,000만배(25Tera / 2Giga) 정도 더 빨라졌다고 볼 수 있다. 대한민국도 슈퍼컴퓨팅 계산자원이 증가함에 따라, 위 <표 3>에서 볼 수 있었던, 1988년 당시에 수행할 수 있었던 시뮬레이션들의

규모 - 자동차 설계 지원 시작, 과학적 기상 예측 등 - 가 현재는 초거대규모의 난류 열유동 해석을 수행해볼 수 있는 규모로까지 성장한 상태이다.

[그림 6] 전 세계 상위 200위 이내의 연산성을 지니는 슈퍼컴퓨터들의 국가별 보유 현황 (2021년 6월 기준)



데이터 출처: <https://www.top500.org/lists/top500/list/2021/06/>

Ⅲ. 연구 패러다임의 변화: GPU, AI, 데이터 기반 계산과학의 보편화

컴퓨팅 하드웨어 및 병렬컴퓨팅 기술의 괄목할만한 성장이 이루어짐에 따라 계산과학적 방법들을 활용하는 연구 트렌드가 점차 보편화되기 시작하였다. 2005년도 전후부터 CPU 성능을 끌어올리는 방식에 대한 패러다임의 변화가 발생하여 CPU의 클럭 자체를 높이는 것보다 CPU가 단순히 단일 코어(single-core)만을 가지는 게 아닌, 점차 더 많은 수의 멀티코어(multi-core)들을 가지도록 하는 쪽으로 CPU들을 개발하여 성능을 더 높이기 시작하였다. 멀티코어 CPU(multi-core CPU)들의 보급은 분산 메모리 병렬프로그래밍(distributed-memory parallel programming)보다 구현이 훨씬 더 용이하고 간편한 공유 메모리 병렬 프로그래밍(shared-memory parallel programming)의 활용을 더욱 촉진하게 된다. CPU 발전 패러다임의 변화와 함께 병렬 프로그래밍의 중요성이 더욱 부각되고 있던 분위기 속에서, 공유 메모리 병렬화 뿐만 아니라 더욱 대규모의 계산을 가능토록 하는 분산 메모리 병렬화 기술들 또한 점차 더 주목받기 시작하였다. 병렬컴퓨팅 기술의 보급 및 이에 기반한 HPC, 슈퍼컴퓨팅 계산 자원들을 활용하는 것이 점점 더 용이해짐에 따라, 계산이 필요한 거의 모든 연구 분야들은 거대 모델들의 시뮬레이션에 도전하기 시작하는 등 연구 범위를 더욱 확장할 수 있게 되었다.

이후 2010년대에 접어들면서 데이터를 기반으로 하여 계산과학이 거의 모든 분야에 응용되기 시작한, 큰 시대적 연구 패러다임(paradigm)의 변화가 발생하였다. 오늘날 대부분의 연구 분야에서 경험하고 있는, 빅데이터(big data)와 인공지능(artificial intelligence, AI)을 기반으로 하는 연구 트렌드의 보편화는 2010년도를 전후로 하여 엔비디아(NVIDIA)를 필두로 그 도화선이 당겨졌다. 그 당시부터 GPU의 연산 능력과 연산 효율은 크게 향상되기 시작하여, GPU 기반의 저비용 연산 효율화의 중요성이 부각되기 시작하였다. 대다수의 계산과학적 연산에 있어서 동일한 비용으로 계산용 컴퓨터를 구성했을 때, CPU 기반보다 GPU 기반의 컴퓨터가 5~10배 이상 연산이 더 빠르기도 하였다. 결국, 계산과학계에서도 GPU 기반의 컴퓨팅을 더욱 활용하는 쪽으로 그 트렌드가 변화하기 시작하였고, 동일한 비용으로 가용할 수 있는 컴퓨팅 리소스가 급증함에 따라 더욱 큰 스케일의 시뮬레이션들도 수행할 수 있게 되었다. 그뿐만 아니라 생명과학 분야 등의 복잡도(complexity)가 상대적으로 큰 시스템들을 다루는 연구 분야들에서도

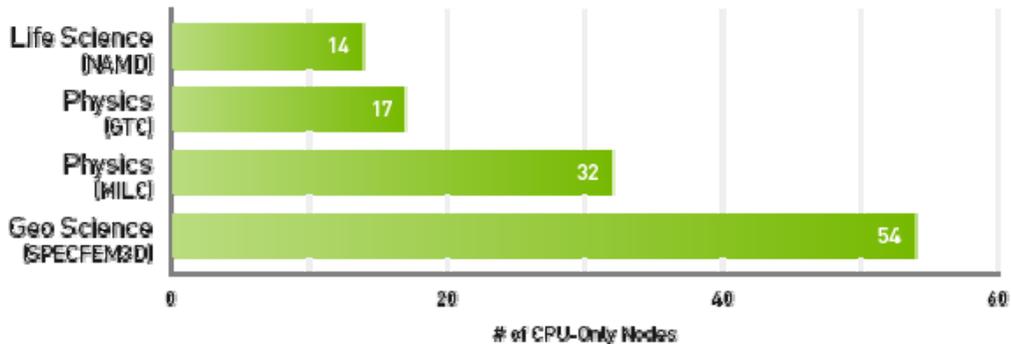
복잡한 난제들에 대해 계산과학적 접근을 시도할 수 있게 됨에 따라, 관련 분야들의 연구 방식에도 자연스럽게 긍정적 변화들이 찾아오게 되었으며, 계산과학이 아우르는 연구 분야의 범위 또한 더욱 넓어지게 되었다. GPU 연산 기술의 보급은 계산이 필요한 거의 모든 연구 분야에 있어 연구 범위를 한 번 더 크게 확장할 수 있도록 하는 긍정적 영향을 주게 되었다. 그런데, GPU 기반으로 크게 도약한 컴퓨터 연산성능의 레벨은 빅데이터를 기반으로 한 인공지능 기술들 개발 및 활용을 더욱 용이토록 하여, 2000년대보다 훨씬 더 크고 혁명적인 연구 패러다임의 변화를 2010년대에 불러오게 되었다.

GPU로 인한 계산과학 분야에서의 전반적인 연산 능력의 향상은 더욱 고도화된 인공지능 모델들의 연구/개발을 촉진시켰을 뿐만 아니라, 인공지능 분야 자체가 계산과학 분야를 이루는 가장 큰 축 중 하나가 되도록 하였다. 또한, 연산 능력의 한계로 그동안에는 접근하기 어려웠던, 데이터 과학(data science)과 인공지능을 기반으로 한 계산과학적 연구 방법들이 학계에 널리 보급될 수 있도록 해주었다. 현재 기준, 동일한 비용을 들여 CPU 기반 또는 GPU 기반으로 딥러닝(deep learning, DL) 연산을 수행하였을 때, 경우에 따라 차이가 있기는 하지만 대략 20배 전후의 연산성능 차이까지 발생하기도 한다. GPU 기반의 과학 컴퓨팅이 수행되지 못했더라면 오늘날의 GPU 하나로도 금방 훈련 시킬 수 있는 옛날에 개발된 깊은 인공신경망(deep neural network) 모델들을 훈련 시키는 것조차 고성능의 병렬컴퓨팅을 통해 훈련 시켜야 했을 것이며, 결국 현재와 같은 수준의 빅데이터 인공지능 기반의 다학제간 계산과학적 연구 패러다임의 변화는 발생하지 않았을 것이다. 계산과학이 다룰 수 있는 영역의 확대뿐만 아니라 AI 기반 컴퓨팅 수준을 비약적으로 향상토록 한, GPU 및 GPU 기반의 병렬컴퓨팅 기술들의 발전은 기존의 계산과학 분야가 하나의 높은 역치(threshold)를 뛰어넘어 한 단계 위로 퀀텀 점프(quantum jump) 해내기 위해 필요했던 핵심 요소였다.

[그림 7] NVIDIA 1 GPU 노드(node)가 14~54 CPU 노드까지 대체하는 사례

1 GPU Node Replaces Up To 54 CPU Nodes

Node Replacement: HPC Mixed Workload



CPU Server: Dual Xeon Gold 6140@2.30GHz, GPU Servers: same CPU server w/ 4x V100 PCIe | CUDA Version: CUDA 9.x | Datasets: NAMD (STMV), GTC (mpi@proc.in), MILC (APEX Medium), SPECFEM3D (four_material_simple_model) | To arrive at CPU node equivalence, we use measured benchmark with up to 8 CPU nodes. Then we use linear scaling to scale beyond 8 nodes.

출처: <https://www.nvidia.com/en-gb/data-center/tesla-v100/>

이제는 GPU 몇 장만 잘 활용하더라도 기계학습(machine learning, ML) 기반의 인공지능 모델을 연구에 활용하는 것이 어렵지 않게 되어, 인공지능 모델들을 학습하기 위한 데이터만 충분히 준비되어 있으면, 별도의 수학적 모델링들을 따로 하지 않더라도 흥미로운 연구 결과들을 도출해 내는 것이 가능하게 되었다. 컴퓨팅 하드웨어 발전 과정에서 일어난 이와 같은 패러다임의 변화는 인공지능 분야의 급속한 발달을 불러옴과 더불어, 거의 모든 연구 분야의 연구 패러다임에 변화가 찾아오도록 만들었을 만큼 대단히 큰 영향을 미쳤다. 그동안 수학적 모델링 등을 통한 계산과학적 접근 자체를 상상조차 하기 어려웠던 인문학, 사회학, 법학, 예술 분야 등과 같은 연구 분야들도 방대한 양의 데이터는 가지고 있었기에, 오랜 기간 축적된 데이터들을 곧바로 인공지능 모델 학습에 활용하여 새로운 의미 있는 연구 결과들을 도출해낼 수 있었다. 이러한 형태의 연구들은 2010년도 이전에는 거의 찾아볼 수 없었던 새로운 연구 형태로, 빅데이터 인공지능 기반의 혁명적 연구 패러다임의 변화는 이전에는 보기 어려웠던 새로운 형태의 계산과학적 연구 결과물들이 쏟아져 나오도록 하였다. 물론, 기존의 계산과학 분야들과 계산과학과 밀접히 접해있던 분야들 또한 인공지능 모델을 활용한 연산 및 시뮬레이션의 극적인 효율화를 달성해 내거나 이전에는 얻지 못했던 새롭고 흥미로운 연구 결과들을 도출해 내는 경우들이 대단히 많아졌다.

IV. 국가별 사례

미국, 중국, 일본, EU는 초고성능컴퓨팅(supercomputing) 경쟁력이 과학과 산업의 미래를 결정할 것이라는 인식하에 정부 주도로 정책역량을 집중하고 있다. 실제로 가용 컴퓨팅 자원의 규모가 커지면 커질수록 수행할 수 있는 연구들의 규모와 가치 또한 비례하여 커지게 되며, 때로는 컴퓨팅 성능 규모가 어느 수준 이상의 문턱을 넘게 되었을 때 얻게 되는 경제적 가치는 컴퓨팅 성능 확충에 투자한 비용을 아주 크게 뛰어넘기도 한다.

미국은 1991년 세계 최초로 초고성능컴퓨팅법을 제정한 바 있다. 그리고 미국은 2016년도에 접어들어 엑사스케일 컴퓨팅 프로젝트를 발족하였으며, 2019년도에는 국가초고성능컴퓨팅 전략을 제정/개정하였다. 일본은 공동 컴퓨팅 인프라 구축 및 자체 기술개발을 위한 Flagship 2020 전략을 수립하였으며(2014~2020), AI 브리징 클라우드 인프라를 2018년도에 도입하였다. 또한 중국은 2016년도에 자체개발, 산업혁신 등을 포함하는 과학기술발전 5개년 계획을 수립한 바 있으며, EU는 2018년도에 초고성능컴퓨팅 생태계 구축을 위해 EuroHPC JU를 결성하였을 뿐만 아니라 유럽 프로세서 이니셔티브를 발족하기도 하였다.

[그림 8] 미·중·일·EU의 초고성능컴퓨팅 경쟁력 향상을 위한 정책역량 집중 사례

<div style="text-align: center;">  <p>세계 최강 유지를 위해 전방위 투자</p> </div> <ul style="list-style-type: none"> ⊙ 세계 최초 초고성능컴퓨팅법 제정('91) ⊙ 국가초고성능컴퓨팅 전략(NSCI) 제정 및 개정('19) ⊙ 엑사스케일 컴퓨팅 프로젝트(ECP) 발족('16) 	<div style="text-align: center;">  <p>인공지능 활용 극대화 등 초고성능컴퓨팅 강국 재도약</p> </div> <ul style="list-style-type: none"> ⊙ 공동인프라 구축, 자체 기술개발을 위한 Flagship 2020전략 수립('14-'20) ⊙ AI 브리징 클라우드 인프라(ABC) 도입('18)
<div style="text-align: center;">  <p>대규모 투자로 초고성능컴퓨팅 강국 부상</p> </div> <ul style="list-style-type: none"> ⊙ 자체개발, 산업혁신 등을 포함하는 과학기술발전 5개년 계획 수립('16) 	<div style="text-align: center;">  <p>독자적 위치 확보를 위해 범 유럽 차원에서 협력</p> </div> <ul style="list-style-type: none"> ⊙ 초고성능컴퓨팅 생태계 구축을 위해 EuroHPC JU 결성('18) ⊙ 유럽 프로세서 이니셔티브(EPI) 발족('18)

출처: 4차 산업혁명 원점점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

또한, 미국, 중국, 일본, EU는 신규 컴퓨팅 인프라 구축에 적극적으로 대규모의 투자들을 해오고 있고, 컴퓨팅 자원을 효율적으로 운용하기 위해 컴퓨팅 자원들을 통합/연동하여 운용하는 공동활용체계를 구축/강화하고 있다. 컴퓨팅 자원 공동활용체계를 효율적으로 잘 운용하면 산업, 연구 분야의 성장 및 국가 전략과제 수행을 강화하는 데 큰 도움이 된다. 예컨대, 미국은 4개 슈퍼컴퓨터센터의 13개 시스템을 묶은 XSEDE를, EU는 TOP 500 슈퍼컴퓨터 성능 순위 50위 내 8개 시스템을 포함한 20개의 컴퓨팅 시스템을 묶은 PRACE, 일본은 9개 대학 3개 국립연구소가 각각 가지고 있는 슈퍼컴퓨팅 자원을 묶은 HPCI라는 컴퓨팅 자원 공동활용체계를 구축하여 활용하고 있다.

[그림 9] 전체 자원량 및 1위 시스템 자원량, 미·중·일·EU 및 한국의 총 자원량



※ 매년 세계 1위국 대비 한국의 자원량 격차 확대: ('11)26배 → ('20)37배

출처: 4차 산업혁명 쿼텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

현재 미국, 중국, 일본, EU는 차세대 엑사스케일 시스템 개발을 위해 자체 기술력을 확보하거나 고도화하는 등 차세대 컴퓨팅 인프라 구축을 위해 기술 경쟁을 하고 있다. 현재의 기술 패권은 세계 최고의 기술력을 지는 민간 글로벌 기업들을 갖춘 미국이 쥐고 있는 상태이고, EU는 기술종속 등을 우려하여 주로 해외에서 도입하여 활용해오던 컴퓨팅 하드웨어들을 자체 개발하는 쪽으로 방향을 전환하기 위해 대규모의 기술 투자를 진행 중이다. 한편, 중국과 일본은 자국 내 전문기업들에 적극적으로 지원하여 슈퍼컴퓨팅 인프라 구축을 위한 자체 기술력을 확보해오고 있다. 특히, 일본의 슈퍼컴퓨터 후가쿠(Fugaku)는 HPL-AI 벤치마크에서 세계 최초로 엑사플롭스(exaFLOPS, Eflop/s) 스케일의 컴퓨팅 성능을 달성한 바 있다.

기록적인 측면에서, 엑사플롭스의 장벽은 코로나19 팬데믹(COVID-19 pandemic)의 과학적 극복을 위해 각광을 받기도 했던 Folding@home 분산 컴퓨팅 프로젝트에 의해 최초로

깨졌고(2020년 3월 1.22 exaFLOPS, 2020년 4월 2.43 exaFLOPS), 단독 슈퍼컴퓨터로는 2020년 6월 일본의 슈퍼컴퓨터 후가쿠(Fugaku)에 의해 엑사 플롭스 스케일의 컴퓨팅이 최초로 달성(1.42 exaFLOPS in HPL-AI benchmark)되었다. 또한, 아래 <표 4>에 있는 2021년 6월 기준 HPL-AI benchmark에 따르면, 전 세계에는 총 2기의 슈퍼컴퓨터(일본 RIKEN의 Fugaku: 2.0 exaFLOPS, 미국 ORNL의 Summit: 1.15 exaFLOPS)가 엑사플롭스 스케일의 컴퓨팅을 달성한 것으로 기록되어 있다.

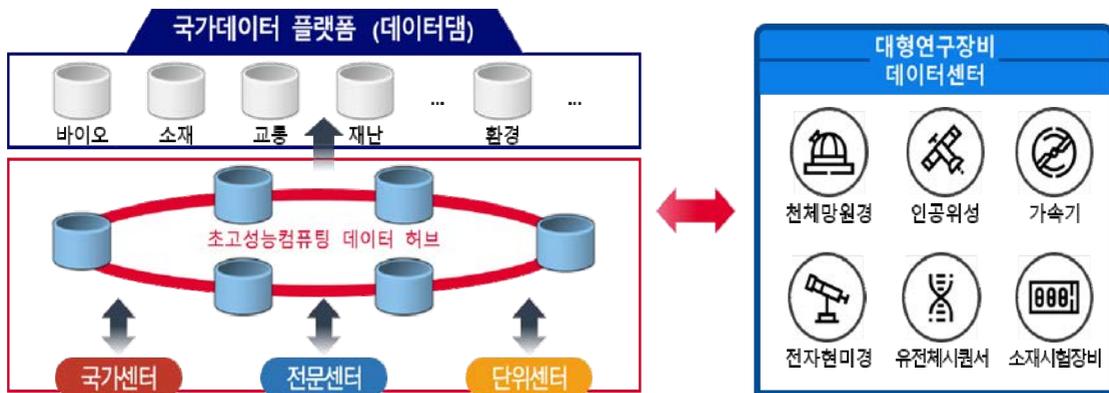
<표 4> HPL-AI Benchmark(2021년 6월)

June 2021							
Rank	Site	Computer	Cores	HPL-AI	TOP500	HPL Rmax	Speedup
				(Eflo/s)	Rank	(Eflo/s)	
1	RIKEN	Fugaku	7,630,848	2.0	1	0.442	4.5
2	ORNL	Summit	2,414,592	1.15	2	0.149	7.7
3	NVIDIA, USA	Selene	555,520	0.63	6	0.063	9.9
4	LBNL	Perlmutter	761,856	0.59	5	0.065	9.1
5	FZJ	JUWELS BM	449,280	0.47	8	0.044	10
6	Univ. Florida	HiPerGator	138,880	0.17	22	0.017	9.9
7	ITC Japan	Wisteria	368,640	0.10	13	0.022	4.5
8	NSC	Berzelius	59,520	0.05	82	0.005	9.9
9	Nagoya	Flow Type I	110,592	0.03	62	0.007	4.3
10	CloutMTS	MTS GROM	19,840	0.015	240	0.002	7
11	ThetaGPU	ANL	280,320	0.012	60	0.007	1.8

출처: <https://icl.bitbucket.io/hpl-ai/results/>

한편, 대한민국은 2011년 초고성능컴퓨터법 제정, 기본계획 수립(2013, 2018) 등을 통해 국가 초고성능컴퓨팅 육성체계를 구축하기 위해 노력해오고 있으며, 한국판 뉴딜 정책에서의 4차 산업혁명을 선도하는 국가로 거듭나기 위한 ICT 산업 분야의 적극적인 예산 지원, 그리고 데이터 댐 구축 프로젝트 등을 통해 미래 계산과학 분야의 성장을 위해 필수적인 컴퓨팅 인프라 구축에 힘쓰고 있다.

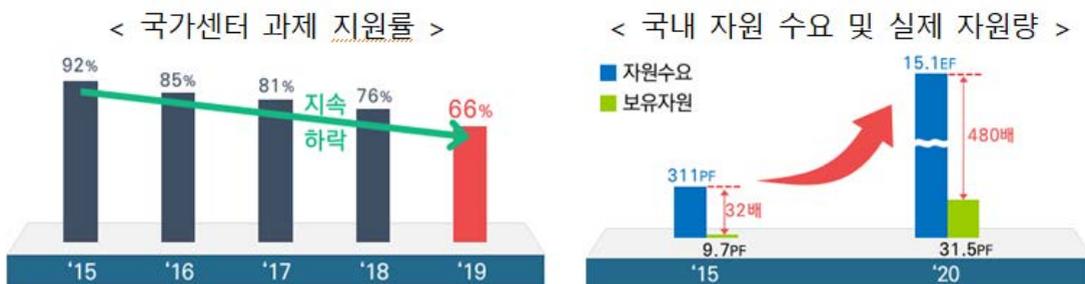
[그림 10] 초고성능컴퓨팅 데이터 허브 체계



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

하지만, 급증하고 있는 초고성능컴퓨팅 활용 수요에 반하여 신규 인프라 투자 정체 및 효율적 활용체계가 미비한 상태로 컴퓨팅 자원의 부족이 심화되고 있는 상황이다. 2021년 6월 TOP 500 슈퍼컴퓨팅 성능 순위 기준, 대한민국 1위 슈퍼컴퓨터의 연산성능은 일본 1위 슈퍼컴퓨터의 연산성능의 약 1/20 수준이며, TOP 500의 상위 100위 이내에 속하는 대한민국의 슈퍼컴퓨터 대수는 일본의 1/6 수준인 3대이다. (일본은 TOP 500 상위 100위 이내 슈퍼컴퓨터 총 18대 보유 중, 2021년 6월 기준)

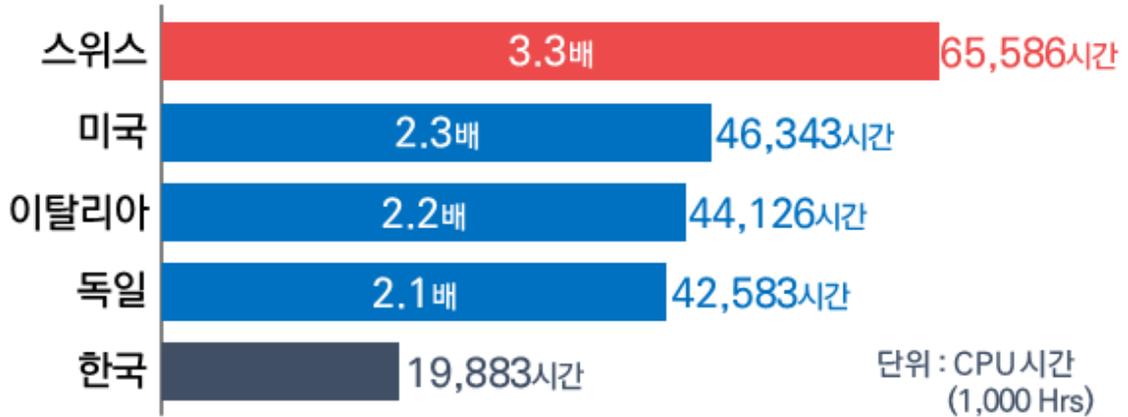
[그림 11] 연도별 국가센터 과제 지원율, 국내 자원 수요 및 실제 자원량



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

그리고 2019년도 기준, 대한민국의 1개 연구당 평균 컴퓨팅 자원 배분 시간은 스위스의 1/3 미만 수준, 미국, 이탈리아, 독일의 1/2 미만 수준이다.

[그림 12] 국가별 1개 연구당 자원 배분 시간(2019)



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

현재 우리나라 정부는 비상경제 중앙대책본부(21-36)가 2021년 5월말에 발표한 “4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략”을 통해 현재 국내 초고성능컴퓨팅센터의 다양화-체계화 및 정부의 적극적인 인프라 구축 지원이 필요하다고 진단하고 있고, 선택과 집중이 분명한 기술개발 로드맵을 만들어 중장기적으로 일관성 있게 추진하는 방향의 필요성과 더불어 산업생태계 구축을 위한 공공부문의 적극적 역할 필요의 필요성, 그리고 우수 연구자-산업계의 혁신적 잠재 수요가 실제 활용으로 이어질 수 있도록 수요 맞춤형 지원 및 인력 양성의 필요성 등에 대해 강조하고 있다.

1 미국

미국은 1991년 세계 최초로 초고성능컴퓨팅법을 제정하였고, 2016년도에는 엑사스케일 컴퓨팅 프로젝트를 발족하였으며, 2019년도에는 국가초고성능컴퓨팅 전략을 제정/개정 하였다. 2020년도 기준 미국은 대략 600~700 PF 규모의 슈퍼컴퓨팅 자원을 갖추었다고 추정되고 있으며, 2021년도 기준 18억 달러 수준의 전방위적인 투자를 바탕으로 2023년도까지 엑사 시스템 3기의 개발을 추진(Exascale Computing Project), 세계 최고의 기술, 산업 패권을 강화하기 위해 노력 중이다. 미국은 과거부터 현재까지 TOP 500 슈퍼컴퓨터 성능 순위 리스트에 압도적으로 가장 많은 횟수로 1위 슈퍼컴퓨터들을 올려오고 있으며, 2020년 6월 일본에 이어,

미국도 2021년 6월에 HPL-AI 벤치마크에서 엑사플롭스(exaFLOPS) 스케일의 컴퓨팅을 달성하였다. (미국 ORNL의 서밋(Summit): 1.15 exaFLOPS)

미국은 연구개발-상용화-산업화로 이어지는 강력한 산업생태계를 갖추고 있고, Intel, IBM, Cray/HPE, NVIDIA 등의 미국 민간 기업들은 컴퓨팅 분야에 있어 글로벌 시장을 주도하고 있다. 막대한 예산을 바탕으로, 미국은 과학 및 연구 분야뿐 아니라 경제와 산업, 의료, 안보 등의 분야들에 있어서도 초고성능컴퓨팅을 통해 혁신 창출이 가능한 다양한 분야들을 국가전략 분야로 선정하여 집중적으로 지원하고 있다. 아래 표는 미국의 초고성능컴퓨팅 중점 활용 분야에 대해 정리하고 있다.

〈표 5〉 주요국의 초고성능컴퓨팅 중점 활용 분야(미국)

국가명	분야	전략(중점) 활용 분야		비고
미국	국가안보	<ul style="list-style-type: none"> 비축관리 핵무기 성능평가를 위한 차세대 시뮬레이션 도구 	<ul style="list-style-type: none"> 적대적 위협환경 및 극초음속 재진입 조건에 대응 	Exascale Computing Project
	과학적 발견	<ul style="list-style-type: none"> 입자물리학 자연의 기본법칙 검증 단백질, 분자구조 분석 	<ul style="list-style-type: none"> 소재탐색, 예측 및 제어 융합플라즈마 모델링 화학원소 기원 파악 	
	경제안보	<ul style="list-style-type: none"> 도시계획 금속부품의 적층 제조 가속화 	<ul style="list-style-type: none"> 지진위험평가 전력망의 안정적이고 효율적 계획 	
	에너지	<ul style="list-style-type: none"> 소형 모듈형 원자로 설계 터빈 출력 발전소 효율 핵융합로 재료 설계 	<ul style="list-style-type: none"> 석유 추출 및 폐기물 처리 고효율, 저공해 연소엔진 설계 스트레스 저항성 직물 분석 	
	헬스케어	<ul style="list-style-type: none"> 약물반응 및 분석 자동화를 위한 예측 모델 구축 	<ul style="list-style-type: none"> 암연구 가속화 	

출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

미국은 현대 계산과학 분야에서의 큰 축으로 거듭난 인공지능(AI) 분야에서도 세계 최고 수준의 기술력을 갖추고 있다. 미국은 구글(Google)과 딥마인드(DeepMind), 애플(Apple), 마이크로소프트(Microsoft, MS), 아마존(Amazon), 아이비엠(IBM), 페이스북(Facebook)과 페어(FAIR), 테슬라(Tesla)와 오픈AI(OpenAI) 등 미국의 글로벌 기업들 중심으로 발전하고 있는 강력한 민간 AI 산업생태계를 갖추고 있다. 딥마인드는 2016년 이세돌 9단을 꺾은

알파고(AlphaGo)를 선보이며 전 세계에 이름을 알리기 시작했고, 최근 알파폴드 2(AlphaFold 2)는 CASP(단백질 구조 예측 능력 평가대회)에서 이례적으로 92점 이상을 기록하는 등 인류의 가장 큰 과학적 문제를 해결하는 데 AI를 핵심 요소로 활용하고 있다. 테슬라는 세계 최고 수준의 자율주행(self-driving) 기술을 보유하고 있고, 테슬라의 CEO 일론 머스크(Elon Musk)가 설립한 오픈AI는 프로그래머와 겨룰 수 있는 AI 개발과 함께(해당 게임명: Dota 2), 현존 최대 규모의 언어모델 GPT-3 및 문장을 이미지화하는 DALL-E를 선보이는 등 세상을 놀라게 하고 있다. 위 미국의 IT 기업들에 의해 뉴립스(NeurlPS)와 ICML 콘퍼런스에 발표된 논문들의 수로도 이들의 세계적 영향력이 얼마나 강한가를 추정할 수 있다. 2020년도 뉴립스에 구글은 총 178편의 논문을 발표, 뒤이어 MS는 95편, 딥마인드 59편, 페이스북 58편, IBM은 38편을 발표하였고, ICML에는 구글은 114편의 논문을 발표, 딥마인드 51편, MS는 49편, 페이스북 34편, IBM 19편, 아마존은 18편을 발표한 바 있다.

2 EU

EU는 2018년도에 초고성능컴퓨팅 생태계 구축을 위해 EuroHPC JU를 결성하였을 뿐만 아니라 유럽 프로세서 이니셔티브를 발족하기도 하였다. 2020년도 기준 EU는 대략 400~500 PF 규모의 슈퍼컴퓨팅 자원을 갖추었다고 추정되고 있다. EU는 2010년대 중반 기술종속 등을 우려, CPU 등 주로 해외 도입을 해오던 컴퓨팅 하드웨어들을 자체 개발하는 방향으로 선회한 상태이며, 2019년도 The European Strategy for High Performance Computing(EC)을 수립, 추진 중이다. 현재 2023년까지 80억 유로 규모의 투자를 통해 자체적으로 엑사시스템을 개발 완료하는 프로젝트 또한 추진 중이다. EU는 별도 벤처회사들의 설립 및 지원을 통하여 슈퍼컴퓨팅 시스템의 구축 및 개발을 위한 공동 R&D, 시스템의 제작/설치 등을 수행하는 형태로 열악한 산업 기반적인 부분들을 극복하기 위한 전략들도 함께 취하고 있다.

EU에 속한 여러 국가들은 오래전부터 슈퍼컴퓨팅에 활용될 수 있는 계산과학 분야의 기반 이론 및 기술들을 개발해낸 역사가 있으며, 기초 학문 영역들에서 경쟁력이 특히 강하다. 물리, 화학, 기계 등의 시스템들을 다루는 여러 대표적인 시뮬레이션 프로그램들은 유럽 국가들에 속한 연구기관이나 기업들에 의해 개발되었다. 아래 표의 주요국의 초고성능컴퓨팅 중점 활용 분야(EU)에서도 EU는 기초과학 분야의 연구들에 보다 초점을 두고 있음을 엿볼 수 있다.

〈표 6〉 주요국의 초고성능컴퓨팅 중점 활용 분야(EU)

국가명	분야	전략(중점) 활용 분야	비고
EU	국가안보	• 사이버 보안 및 범죄에서 특히 중요한 인프라의 보호를 위해 활용	The European Strategy for High Performance Computing (EC, 2019)
	과학	• 기초물리학: 물질, 우주에 대한 이해 • 물질과학: 제약, 에너지 분야의 새로운 중요 요소 설계 • 지구과학: 행성 수준의 대기 및 해양 현상 모델링	
	산업	• 자동차, 항공우주, 재생에너지, 건강 등을 혁신하고, 생산성을 높이며, 더 높은 가치의 제품과 서비스 창출에 활용	

출처: 4차 산업혁명 권תר점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

3 중국

중국은 2016년도에 자체개발, 산업혁신 등을 포함하는 과학기술발전 5개년 계획을 수립한 바 있다. 2020년도 기준 중국은 대략 500~600 PF 규모의 슈퍼컴퓨팅 자원을 갖추었다고 추정되고 있다. 중국은 11차 5개년계획('06~'10) 기간, 국가 863계획을 통해 슈퍼컴퓨터 개발 본격화하였고, 대규모 내수시장을 기반으로 기업과 협업하여 기술력 확보 및 시장 진입 추진을 위해 힘쓰고 있다. 중국은 약 40여 년 전부터 슈퍼컴퓨터 개발에 집중적으로 투자해왔고, 다우닝(Dawning) 등의 마이크로프로세서 개발 전문기업들을 육성, 슈퍼컴퓨터에 들어갈 CPU 칩 등의 설계, 제조에 대한 원천기술들을 확보하고 있다.

이미 중국은 과거 세계 최고 슈퍼컴퓨팅 시스템을 여러 차례 구축한 이력들이 있다. 2010년 텐허(天河)-1A, 2013년 텐허(天河)-2, 그리고 2017년 쑤웨이 타이후 라이트(Sunway-TaihuLight)에 의해 기록된 TOP 500 슈퍼컴퓨터 성능 순위 1위 기록들은 과거 중국에 의한 기록들이며, 아래 표에서는 쑤웨이 타이후 라이트의 대표 응용 사례들을 살펴볼 수 있다. 현재 중국은 일본과 미국의 기술력에 대응하기 위해 연간 1조 원 이상을 투자하여 엑사 시스템 개발/구축 등을 위한 기술개발에 박차를 가하고 있다.

〈표 7〉 중국의 슈퍼컴퓨터 쑤웨이 타이후 라이트의 대표 응용 사례들

Table 4 Summary of the major applications on the Sunway TaihuLight, compared with similar applications on other large-scale systems

Category	System	Application summary	Scale of run	Performance
Non-linear solver	Sunway TaihuLight	A fully-implicit nonhydrostatic dynamic for cloud-resolving atmospheric simulation	131072 MPEs and 8388608 CPEs	1.5 PFlops
	Sequoia	An implicit solver for complex PDEs in highly heterogeneous flow in Earth's mantle [3]	1572864 cores	687 TFlops
Molecular dynamics	Sunway TaihuLight	Atomic simulation of silicon nanowires	131072 MPEs and 8388608 CPEs	14.7 PFlops
	Tianhe-1A	Molecular dynamics simulation of crystalline silicon [20]	7168 GPUs (3211264 CUDA cores)	1.87 PFlops
Phase-field simulation	Sunway TaihuLight	Coarsening dynamics based on Cahn-Hilliard equation with degenerated mobility	131072 MPEs and 8388608 CPEs	39.678 PFlops
	Tsubame 2.0	Dendritic solidification [6]	16000 CPU cores and 4000 GPUs (1792000 CUDA cores)	1.017 PFlops

출처: <http://scis.scichina.com/en/2016/072001.pdf>

또한, 중국은 현대 계산과학 분야에서의 큰 축인 인공지능 분야에서도 세계 최고 수준의 기술력을 갖추고 있는 것으로 평가되고 있다. 최근 중국은 AI 학술 논문 발표량 부문과 AI 특허 출원량 부문 모두에서 세계 1위를 달성, AI 연구의 양적인 부분에서 세계 최고 수준을 해마다 계속 유지해오고 있다. 그리고 최근 10년간 ACM과 AAAI에 발표된 논문들의 수에서 칭화대가 세계 1위(2019)를 차지하기도 하는 등 AI 연구의 질적인 부분에서도 중국은 세계 최고 수준을 달리고 있는 것으로 평가되고 있다. 중국은 2004년부터 현재까지 2016년을 제외하고 AI 논문 출판 수는 항상 미국을 앞서왔으며, AI 인덱스 리포트 2021에서도 2020년도 중국의 AI 논문 피인용 수는 미국을 제치고 1위를 달성한 바 있다. 중국은 미국 다음으로 AI 인재를 많이 보유하고 있다고 알려져 있으며, 정부 주도로 AI 기술 및 인재들의 육성이 이루어지고 있다. 막대한 데이터의 보유와 함께, 중국의 바이두와 텐센트 같은 민간 기업들도 세계 최고 수준의 AI 기술들을 갖추고 있으며 빠른 속도로 성장하고 있다고 추정되고 있지만, 이 기업들의 기술 수준과 성장 속도의 정도를 가늠하기가 상당히 어렵다는 시각들이 있다.

4 일본

일본은 공동 컴퓨팅 인프라 구축 및 자체 기술개발을 위한 Flagship 2020 전략을 수립하였다(2014~2020). 2020년도 기준 일본은 대략 600 PF 내외 규모의 슈퍼컴퓨팅 자원을 갖추었다고 추정되고 있다. 일본은 자국 내 초고성능컴퓨터 전문기업(NEC, Fujitsu)을 적극적으로 지원하는 정책을 바탕으로 컴퓨팅 하드웨어의 국산화 추진해오고 있다. 일본 기업 후지쯔(Fujitsu)는 64비트 ARM 아키텍처 기반(저전력, 고효율에 초점)의 마이크로프로세서인 A64FX 48C 2.2GHz를 개발하여 현재 세계 최고 성능을 지니는 슈퍼컴퓨터인 일본 리켄(RIKEN) 후가쿠(Fugaku)에 탑재하였다.

ARM 아키텍처 기반으로 세계 최고 성능을 달성한 일본의 접근은 현재 주로 GPU 기반으로 최고 성능을 뽑아내고 있는 트렌드에서 벗어나 있어 차별적이다. A64FX는 512비트 벡터 구현이 포함된 ARMv8.2-A Scalable Vector Extension SIMD(single instruction multiple data) 명령어 세트를 사용하는 최초의 프로세서라는 특징이 있다. 영국의 회사였던 ARM은 2016년 손정의 회장이 이끄는 일본의 소프트뱅크(SoftBank)에 의해 인수되었다. (일본의 소프트뱅크는 35조 원 가량을 들여 ARM의 주식들을 전액 현금으로 전량 매입한 바 있다.) ARM은 스마트폰의 두뇌인 모바일 애플리케이션 프로세서(application processor, AP)의 95% 가량을 설계하고 있는 회사로, 일본은 현재 민간 기업들 차원에서 연산 처리장치 설계 및 제조에 대한 원천기술들을 확보하고 있다.

일본은 자연재해라는 현실적 문제 해결을 위해 '81년부터 기술개발 시작, '11년(K-Computer, 1조 3천억 원), '20년(Fugaku, 1조 5천억 원) 세계 최고 시스템을 개발한 바 있고, 초스마트 사회(Society 5.0)와 같은 정책을 추진하고 있다. 일본은 현재 세계 1위 성능을 지니는 슈퍼컴퓨터인 후가쿠(Fugaku)를 보유하고 있으며(TOP 500 슈퍼컴퓨터 성능 순위 기준), HPL-AI 벤치마크에서도 2020년도 6월 단일 슈퍼컴퓨터로는 세계 최초로 엑사플롭스(exaFLOPS) 스케일의 컴퓨팅 성능을 달성하였다. (2020년 6월 1.42 exaFLOPS, 2020년 11월~2021년 6월 2.0 exaFLOPS)

최근 일본은 자국의 대표 국책 연구기관인 신에너지·산업기술종합개발기구(NEDO)를 통해 인공지능(AI) 신기술 개발을 위한 액션 플랜(action plan)을 발표하면서 AI를 적극적으로 활용할 수 있는 제조·도시·수송·교육·건강 분야 중심으로 아래 <표 8>과 같이 12가지 중점 기술들을

선정하였다. 또한, 아래 <표 9>에서 기술된 전략(중점) 활용 분야들에서 슈퍼컴퓨팅을 활발히 활용하고 있다.

<표 8> 인공지능(AI) 신기술 개발을 위한 액션 플랜(action plan)

인공지능(AI) 신기술 개발을 위한 일본 NEDO의 액션 플랜	
의미 이해 AI	현재 대규모 언어모델들의 숨어있는 의미를 이해하지 못하는 한계 극복, 뇌 활동의 주요 속성인 메타인지 기법을 갖춘 AI 개발 필요. 기호 추론과 심층학습을 결합할 수 있는 기술적 돌파 등이 필요.
심층 강화 학습	현재 AI는 서로 대화할 수 있는 기능 취약. 일반 언어로 세상을 연결하는 모델 생성 필요. 기존 AI가 할 수 없는 언어처리를 가능토록.
시뮬레이션 X 기계학습	연역적 사고와 데이터에 기반 추론적 사고를 조합하여 예측, 추론.
메타인지 AI 개발	인간 뇌의 주요 활동인 메타인지(자신의 인지 과정을 객관적으로 보는 능력)를 AI로 연결하는 기술.
전이학습기술	목표로 하는 대상에 대해 실험을 많이 하지 못하더라도 다른 유사한 대상에 관한 학습을 통해 얻은 지식을 학습에 활용.
제조프로세스 최적화 AI	개발 및 제조 공정 전반을 자율적으로 최적화, 운용.
다품종소량생산 AI	소수의 직원이 여러 공정을 책임지고 완제품을 만들어내는 것과 같은 프로세스들에 AI 도입, 프로세스 효율화.
분자설계 AI	최신 AI 모델들을 개발/활용해 분자설계, 연구/산업 분야에서의 활용.
무인운반(AGV)용 환경인식 기술	다양한 환경적 요소들을 동시에 인식/활용, 어떤 환경에서도 안전하게 활동할 수 있는 AGV 개발.
AI에 의한 학습지원	인간이 학습하는 과정을 AI에 의해 밝히는 기술이 필요. 인간이 아직 알지 못하는 효율적인 학습법 탐색. AI를 통한 학습지원 개인화.
AI 주치의	AI의 인간 진단 일상화. 자연언어 이해 기술, 진단/판정 기술들의 개발 및 통합적 활용.
휴먼 디지털 트윈	개인의 데이터나 생활 행동 습관 등을 파악할 수 있는 각종 데이터로 디지털 트윈 생성, 의료 분야에 응용.

출처: 일본 신에너지·산업기술종합개발기구 NEDO(2021.06.)

〈표 9〉 주요국의 초고성능컴퓨팅 중점 활용 분야(일본)

국가명	분야	전략(중점) 활용 분야	비고
일본	의료	• 예측, 예방, 개인맞춤형, 참여 확대 중심 건강수명 연장	Society 5.0
	이동혁명	• 사람·물류의 이동 무인화	
	산업	• 제조프로세스(개발-제조-판매-소비) 개선	
	건설	• 건설현장 생산성 향상	
	금융	• 데이터 유통환경, 비현금사회, 전자결제 보안 강화 등 핀테크 환경 조성	

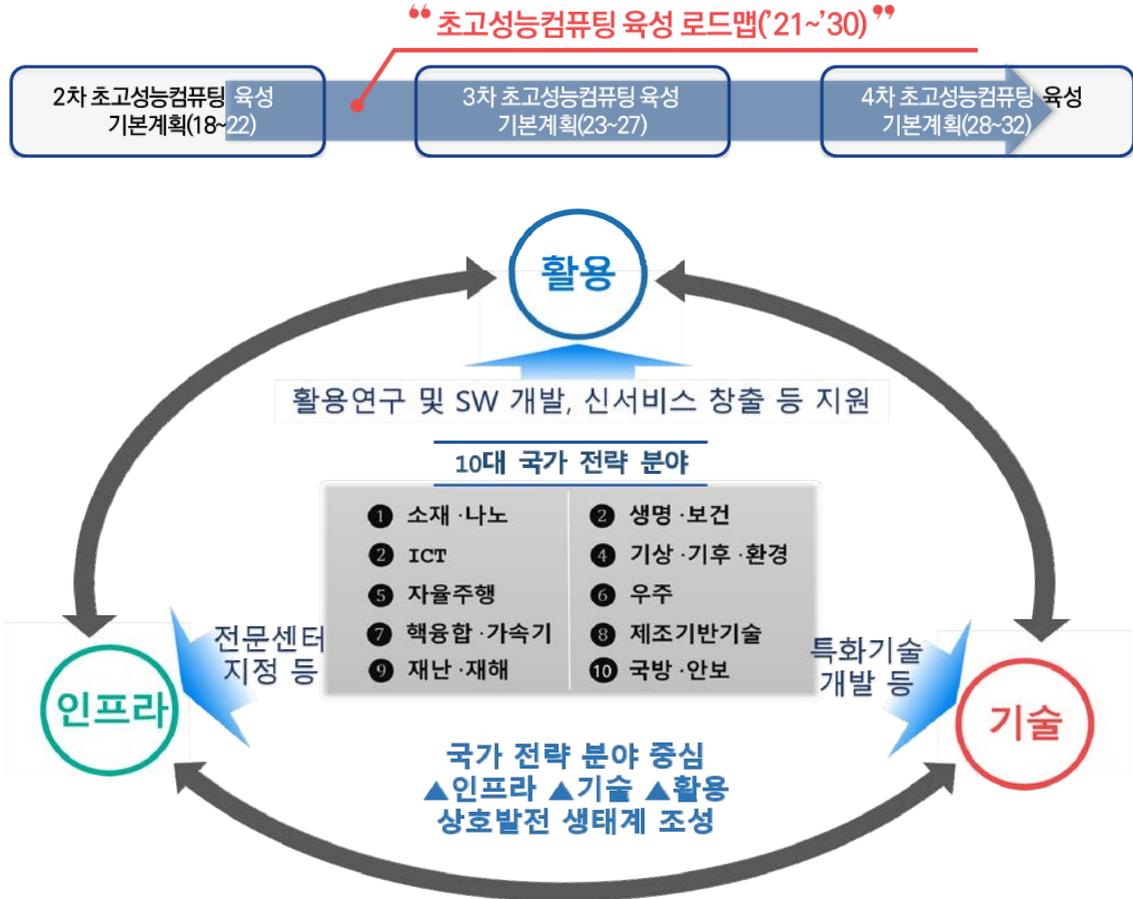
출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

5 대한민국

대한민국은 1988년 슈퍼컴퓨터 1호기 도입을 시작으로, 가장 최근 2018년도에 5호기가 도입되어 거대 계산과학 프로젝트들에 활용되고 있다. 대한민국 1위 성능을 지니는 슈퍼컴퓨터는 초고성능컴퓨터 TOP 500 순위 내 대략 10~40위 사이에 위치해오고 있었고, 국내 초고성능컴퓨팅 분야 정부 투자액은 연평균 1,000억 원 내외 수준으로, 주요국 대비 1/10 수준인 상황이다. 또한, 해외 주요국의 슈퍼컴퓨팅 시스템의 평균 운영 기간은 대략 4~5년인 반면, 대한민국의 경우 슈퍼컴퓨팅 시스템의 평균적인 운영 기간은 대략 8년 내외 수준이었다.

2020년도 기준 대한민국은 대략 20 PF 이내 규모의 슈퍼컴퓨팅 자원을 갖추었다고 추정되고 있다. 그리고 이는 일본의 약 1/30 수준이다. 현재 우리 정부 관계 부처들은 페타급 슈퍼컴퓨팅 기술에서 엑사급 기술로의 대전환기 속에서 선제적 대응을 통한 선도국과의 격차를 조기 극복, 신성장 기회를 창출할 필요성을 강조하고 있고, 2018년도에 수립된 2차 기본계획 이후 국내의 패러다임 변화에 대응한 장기 전략 마련의 시급성을 인지하고 있다. 2021년 5월 28일 비상경제 중앙대책본부(21-36)의 “4차 산업혁명 퀀텀 점프를 위한 국가초고성능컴퓨팅 혁신전략”에서는 아래 그림과 같은 흐름의 2030년도까지의 초고성능컴퓨팅 육성 로드맵을 제시하고 있다.

[그림 13] 초고성능컴퓨팅 육성 로드맵 및 추진 전략 도식화(2021~2030)



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

위 정부 로드맵에서의 중점 추진 과제들은 크게 1) 전략적 인프라 확충, 2) 독자적 기술 확보 및 산업화 기반 마련, 3) 혁신적 활용 활성화로 나누어진다. 가장 먼저 전략적 인프라 확충을 위해, 세계적 인프라를 갖춘 국가센터를 육성할 계획이 세워져 있고, 분야별 전문센터를 지정, 육성하는 부분과 초고성능컴퓨팅자원 공동활용체계 2.0 구축이 추진될 예정이다. 한편, 독자적 기술 확보 및 산업화 기반 마련을 위해서는 전략기술 포트폴리오 기반의 핵심원천기술 확보와 자체 프로세서 기반 역사급 초고성능컴퓨터 개발 및 구축과 더불어 기술사업화 장벽 해소 및 지속성장기반 구축을 위한 과제들이 추진될 예정이다. 마지막으로, 혁신적 활용 활성화를 달성하기 위해 국가 전략 분야 중심의 수요 맞춤형 지원을 강화할 예정과 함께, 전문성 기반의 개방형 활용 생태계 구축이 추진될 예정이다.

정부는 다음과 그림과 같이 14개 후보 분야 중 초고성능컴퓨팅 활용 파급효과가 큰 10개 국가 전략 분야들을 선정하여 해당 분야들을 중심으로 컴퓨팅 자원을 배분하고, 연구계와 산업계, 공공서비스 각 분야에 대해 수요 맞춤형으로 지원할 계획을 갖추고 있으며, SW 생태계 구축, 연구개발서비스업, 전문인력 육성 등을 통해 초고성능컴퓨팅의 지속적인 발전 기반을 구축할 계획을 가지고 있다.

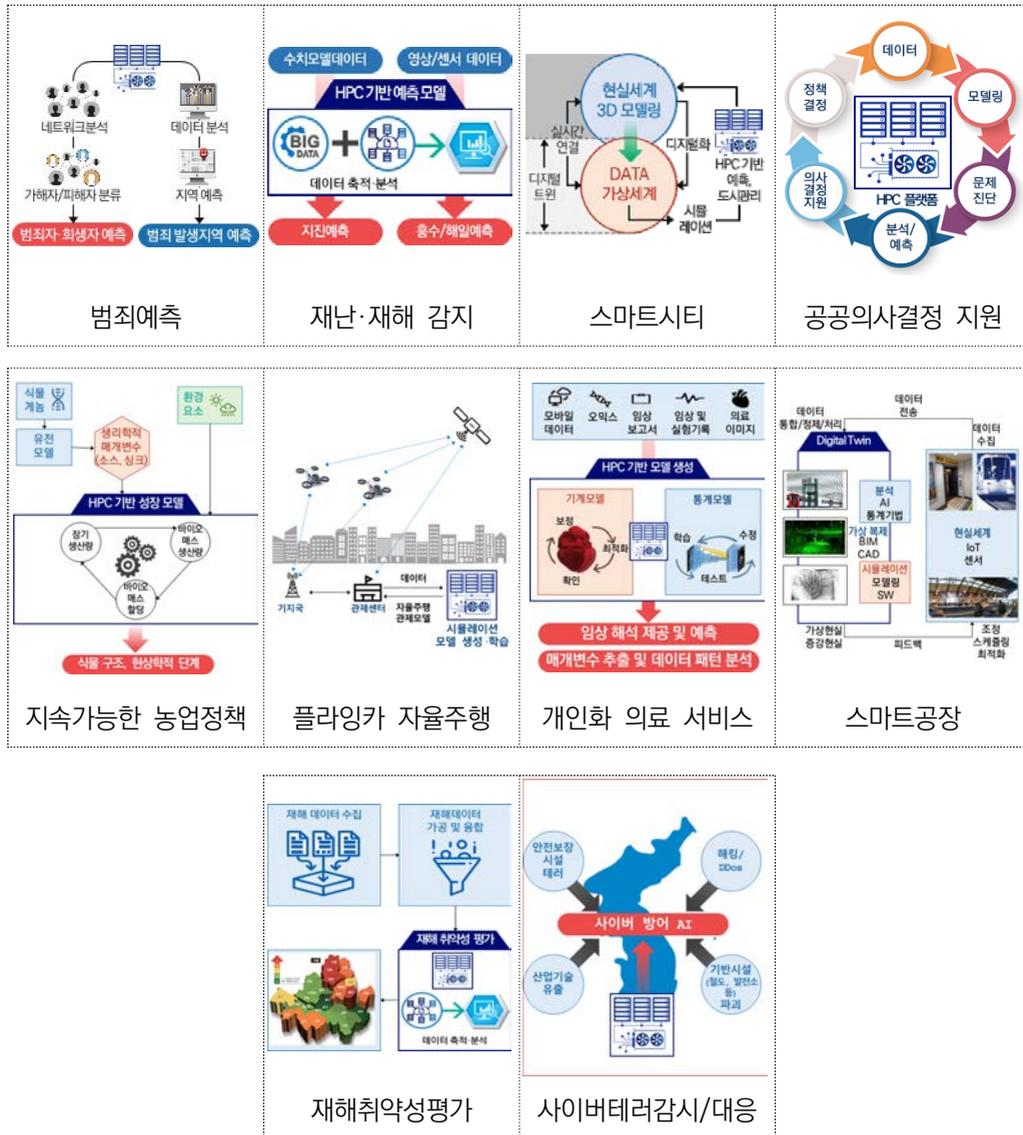
[그림 14] 10대 전략분야별 유망기술 예시



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

그리고 더 나아가, 초고성능컴퓨팅 기반의 혁신적 서비스 모델을 발굴하고, 애플리케이션 개발 및 초기 운영을 위한 전용 자원 제공 등의 지원 확대(30년까지 新서비스 10개 창출)를 통해, 국가 전략 분야 공공-비즈니스 新서비스 창출을 위한 계획들이 추진될 예정이다.

[그림 15] 초고성능컴퓨팅 기반의 新서비스 예시



출처: 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(비상경제 중앙대책본부 21-36, 2021)

대한민국은 최근 “4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(2021. 05. 28., 비상경제 중앙대책본부 21-36)”에 기술된 전략들을 기반으로 추진 예정인 정부 로드맵에 따라, 향후 정부는 ‘범부처 국가초고성능컴퓨팅 협의체’ 구성·운영을 통해 관련 사업 기획 및 예산 확보, 성과 점검 등을 체계적으로 추진(‘21년~)할 계획으로, 同 혁신전략은 기본계획과 연계, 주기적으로 현행화하여 이행력 강화할 예정이며, 예비타당성조사가 필요한 사항(국가 플래그십

초고성능컴퓨터 구축·운영, 범부처 국가초고성능컴퓨팅자원 선진화, 초고성능컴퓨팅 원천기술 개발 및 시스템 실증 등)은 예타 절차를 거쳐 추진 여부를 최종 결정할 예정이다. (국가초고성능컴퓨팅 혁신전략(2021. 05. 28.))

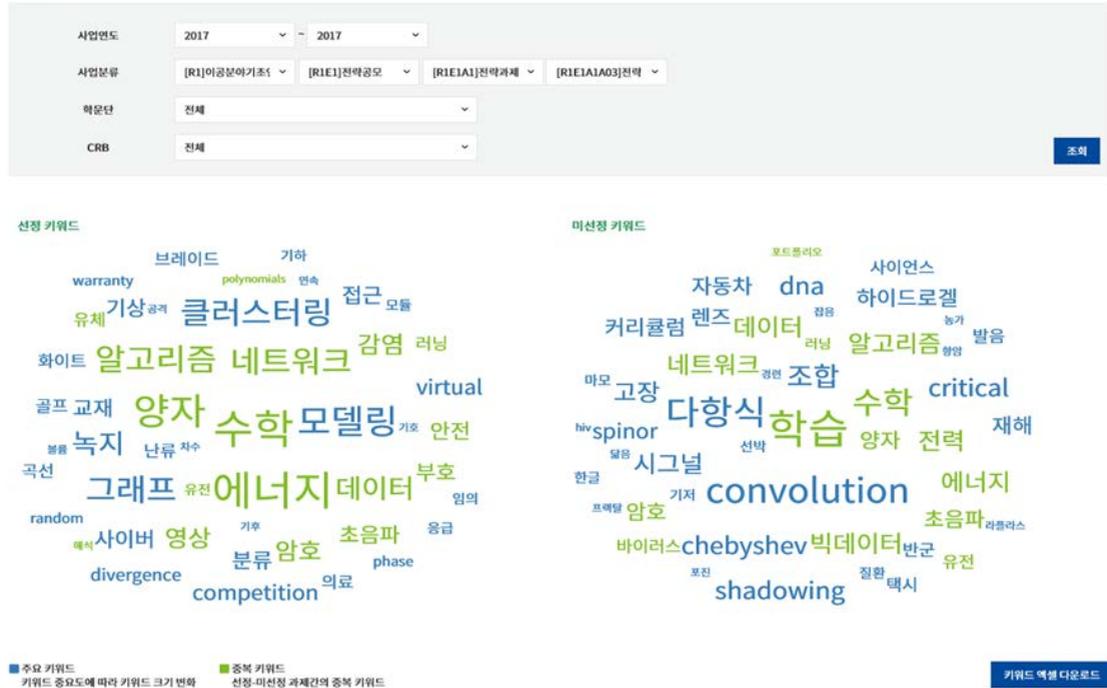
한국연구재단에서도 2017년 기초연구사업 전략과제로 산업수학 프로젝트를 시작하여 마지막 해 지원을 하고 있다. 사업 목표는 “4차 산업혁명시대의 근간이 되는 수학의 기초 원리·이론 연구 및 수학 이론의 개발·적용·응용 연구 지원을 통한 미래기술 확보·인력양성·생태계 확충 등 4차 산업혁명 시대 준비 및 관련 분야 기초연구 역량 강화”이다. 지원유형은 4차 산업혁명 시대에 필요한 기술 및 혁신 활동에서 수학적 방법론과 접근이 중심적인 연구를 대상으로 한다. (서진근 외, 2017)

- ① 혁신 기술 관련 수학적 원리와 이론
- ② 신산업, 새로운 형태의 국민 생활 등 관련 수학 기반 응용
- ③ 수학을 활용한 기존 산업 활동 향상
- ④ 기타 분야와의 융합 또는 인력양성

아래 빅데이터 분석결과는 2017년 시작되어 5년차에 진입하게 되는 전략과제였던 “산업수학”의 선정과제와 미선정과제의 키워드 클라우드이다. 빈도를 기반으로 작성되었으며, 한국연구재단 지능형 연구지원 분석 포털(트렌드분석시스템)을 통해 도출한 결과이다. (양정모, 2021)

[그림 16] “산업수학”의 선정과제와 미선정과제의 키워드 클라우드

트렌드 분석



키워드에서 볼 수 있듯이 산업 수학의 확산 시점에 선정된 많은 과제들이 계산과학을 중심으로 응용수학 영역에서 선정되었다. 영문 키워드 네트워크 분석 클라우드¹⁾에서 볼 수 있듯이 부록에서 소개하고 있는 계산과학의 개념과 연결되는 키워드를 찾아볼 수 있다.

1) 모델링 : 중요도(TF-IDF, Text Frequency), Window size : 3

V. 맺음말

1980년대 전후로 본격적으로 자리를 잡아가기 시작한 계산과학 분야의 탄생에는 20세기 중반 컴퓨터의 등장과 더불어, 그 이전부터 이론적으로 준비되어 온, 계산과학 분야에 활용될 수 있는 수학 분야들과 수학적 도구들 및 방법론들의 발전 역사가 필연적으로 앞서게 된다. 그리고, 컴퓨팅 하드웨어의 지속적인 발달에 따른 가용 컴퓨팅 자원이 증가하는 정도에 따라, 중요한 하드웨어 발전에 대한 패러다임 변화가 발생할 때마다 계산과학 분야가 아우르는 범위가 그만큼 더 확장되어 왔다. 특히, 2010년대쯤부터 GPU 기반의 병렬컴퓨팅(GPU-based parallel computing)이 점진적으로 보편화되기 시작하여 가용 컴퓨팅 자원이 큰 폭으로 증가하기 시작하였다. 이에 힘입어, 인공지능 겨울(AI winter) 시기부터 제한된 계산자원 등으로 힘겹게 발전해 왔던 인공지능 분야의 괄목할만한 성장이 부각됨과 함께, 빅데이터를 기반으로 한 인공지능 기술이 널리 보급되기 시작하였다. 빅데이터 기반 인공지능 기술 활용의 보편화는 딥러닝(deep learning)을 포함한 기계학습(machine learning) 분야가 여러 연구 분야들에 쌓여있던 빅데이터(big data)와 만나도록 해주었다. 이 시기부터 거의 모든 분야에서 인공지능이 활용된 수많은 연구 결과들이 쏟아져 나오며 따라, 인공지능 기술을 매개로 한 다학제간 극적인(dramatic) 연구 시너지(synergy)들이 연쇄적으로 일어났다. 계산과학 분야를 포함한 전반적인 연구 분야들의 연구 패러다임(paradigm) 자체에 커다란 시대적 변화들이 찾아오게 되었다.

현대 과학에서 계산과학 분야가 이론 및 실험과학 분야와 더불어 하나의 새로운 축을 형성하기 시작한 이후, 계산과학이 각종 연구에 활용되는 비중은 시간이 지날수록 모든 과학과 공학 분야에서 커져 왔으며, 거의 모든 과학, 공학 분야에 있어서 계산과학 기술은 공통적인 기초 학문이 되었다. 전 세계의 수많은 대학들은 여러 형태의 계산과학 협동과정들을 개설하여 계산과학과 융합된 형태의 다학제간 연구를 촉진시키기 위한 기반을 다지는 데 큰 힘을 쏟고 있다. 특히, 국가 차원의 연구 경쟁력을 강화하는 차원에서, 전 세계의 각국은 가용 슈퍼컴퓨팅 자원을 최대한 많이 확보하기 위해 치열한 경쟁을 하고 있는 추세로, 이는 현대 과학에서 있어서 계산과학이 차지하는 비중과 중요성이 어느 정도로 거대한지를 가늠할 수 있게 한다. 2020년 6월 HPL-AI

벤치마크에서 엑사플롭스(exaFLOPS) 스케일의 컴퓨팅을 최초로 달성했던 일본을 선두로 미국, EU, 중국 등은 엑사스케일(exascale) 슈퍼컴퓨터를 확충하기 위해 치열한 노력을 기울이고 있다. 한편, 2021년 6월 기준 대한민국 1위 슈퍼컴퓨터의 연산성능은 일본 1위 슈퍼컴퓨터에의 연산성능 비해 대략 20배 이상 뒤쳐져있는 상태이다.

현재 국가 차원에서 확보된 슈퍼컴퓨팅 자원들이 일본, 중국 등의 주변 아시아 국가들에 비해 크게 부족한 실정이긴 하지만, 우리나라도 병렬컴퓨팅 및 슈퍼컴퓨팅의 중요성을 2010년도 대 이전부터 미리 알고 있었기에 2011년 초고성능컴퓨터법 제정 및 기본계획 수립(2013, 2018)과 더불어, KISTI를 중심으로 병렬프로그래밍 교육을 국가 차원에서 지속해서 추진해오고 있으며, 컴퓨팅 자원이 부족한 연구자들에게 슈퍼컴퓨팅 자원을 효과적으로 지원해 줄 수 있도록 오랜 기간 노력해오고 있다. 연구 분야들뿐만 아니라 거의 모든 산업 분야들도 계산과학적 접근을 통한 비용 및 시간 절감효과를 크게 달성할 수 있음을 인지하고 있다. 주로 대기업들을 중심으로, 다수의 국내 기업들은 HPC 자원 확보 및 CAE(computer-aided engineering) 인력 확보와 역량 강화에 큰 노력을 기울이고 있다. 반면, 중견/중소기업들은 동일 업종 미국, 유럽 등의 중견/중소기업들에 비해 계산과학적 방법들의 활용 수준이 크게 뒤쳐져있다고 알려져 있어, 계산과학 기술 역량의 신장을 통한 중견/중소기업들의 성장 여력이 크게 열려있는 상황이다. 따라서, 정부 차원의 계산과학 교육, 컴퓨팅 인프라 등에 대한 지원 강화, 정부와 민간 차원의 협력 강화, 산학연 협력 강화 등에 대한 중장기적인 노력을 보다 적극적으로 기울이면, 앞으로 독립된 분야들 사이에서의 첨단계산과학 지식의 연결 시너지를 통한 전반적인 연구/산업 분야들의 상호 발전 효과까지도 기대해 볼 수 있을 것이다. 현재, 아주 넓은 분야들에 걸쳐 서로 흩어져 있는 계산과학적 지식의 종류들의 수와 규모는 상당한 수준이다.

이러한 큰 시대적 흐름은 분야를 뛰어넘어 누구나 첨단계산과학에 필요한 기초 수학, 수치해석 및 수학적 방법론들, 알고리즘, 데이터 과학, 병렬컴퓨팅, 모델링, 시뮬레이션, 시각화 기법 등의 제반 기본 지식에 대한 교육을 쉽게 받을 수 있도록 하는 국가 차원의 교육 인프라 구축 및 개선을 요구하고 있다. 동시에 각종 과학, 응용과학, 공학 등의 분야에서 핵심적으로 활용되고 있는 계산과학적 방법들에 대해 필요할 때 언제든지 교육, 학습하여 응용할 수 있도록 하는 기회가 제공될 수 있도록 하는 교육 인프라 구축도 필요하다. 충분한 수준의 교육 인프라 구축만으로도 전반적인 연구, 산업 분야들의 국가 경쟁력을 크게 신장할 수 있을 것이다. 물론, 대한민국의 슈퍼컴퓨팅 자원의 규모를 일본, 중국 등의 주변 국가들과도 어느 정도 견줄 수 있는 수준으로 크게 끌어올려, 과학적 계산들을 위한 가용 컴퓨팅 자원의 규모를 충분히 키워줌과 동시에 연구자들이 손쉽게

부족하지 않게 슈퍼컴퓨팅 자원에 접근할 수 있도록 컴퓨팅 인프라를 충분히 확충토록 하는 일련의 작업들 또한 국가 경쟁력 향상을 위해 필요하다.

고성능 병렬컴퓨팅, 그리고 빅데이터 기반의 인공지능 기술들의 보편화와 함께 그 중요성이 훨씬 더 높아진 현대 계산과학은 단순히 하나의 독립된 연구 분야의 성격을 넘어, 거의 모든 분야에 활용될 수 있게 된 하나의 기초 학문과 같은 성격으로서 자리매김하고 있다. 현시대 속에서 일어나고 있는 거대한 연구 패러다임의 변화에 대해, 우리는 이러한 현재의 시대적 흐름을 무게감 있게 인지해나가며 미래를 철저히 준비해나갈 필요가 있다.

VI. 부록

1 계산과학의 정의

계산과학(computational science)은 수치적 방법 및 고급 컴퓨팅 기술(advanced computing technique)들을 사용하여 복잡한 문제들을 이해하고 해결하는 빠르게 성장하고 있는 분야이다. 계산과학은 수많은 분야들을 아우르고 있으며, 자연계 및 여러 공학적 시스템 등의 계산/수학적 모델링(mathematical modelling), 시뮬레이션 이론 및 관련 기술들의 개발을 모두 포함하고 있다. 여기에서 계산에 필요한 제반 이론과 기술들은 계산과학기술 또는 과학적 계산(scientific computing)이라 일컫기도 한다. 예를 들어, 유한요소법(finite element method, FEM), 몬테카를로 방법(Monte Carlo method)은 대표적인 계산과학기술들이다.

계산과학 분야에서는 물리/화학적, 생물학적 문제뿐만 아니라 공학, 사회학, 인문학적 문제들을 해결하기 위해 다양한 알고리즘들을 개발한다. 또한, 거대한 계산자원 사용을 요구하는 문제들의 해결을 위해 필요한 고급 시스템 하드웨어, 펌웨어, 네트워킹 및 데이터 관리 구성요소들을 개발하고 최적화하는 부분, 그리고 컴퓨팅 인프라 자체를 효과적으로 구축, 개발, 관리, 지원하는 부분들도 계산과학의 영역에 속해있다. 계산과학 분야에서의 과학적 연구 방식은 컴퓨터를 이용한 다양한 계산방식들을 통해 현상을 예측, 이해하는 방식을 취하며, 주로 이론과 실험 두 가지 방식으로만 수행되던 기존의 과학에 컴퓨터에 구현된 수학적 모델의 분석, 계산, 시뮬레이션 등을 통해 기존의 이론 및 실험을 통한 연구를 보완하거나 연구 속 한계점들의 극복을 돕는다. 이를 위해 다양한 매개 변수 들을 입력받아 실행되는 계산과학적 프로그램 및 소프트웨어들이 개발되며, 다양한 수치 알고리즘 및 계산 수학 이론 등이 함께 활용된다. 경우에 따라, 계산과학 프로그램들은 슈퍼컴퓨팅을 활용한 거대 계산을 수행하기 위해 사용되기도 한다. 그리고 시간이 흐름에 따라 슈퍼컴퓨팅 플랫폼 구축 및 활용의 중요성은 더욱더 강조되는 추세이다. 계산과학 자체 또한 시간이 지날수록 인공지능 기술들이 융합된 형태로 그 중요성이 더욱 강조되고 있는 상황이다.

현대 과학에서 계산과학은 과학 분야 내에서의 큰 한 축으로 간주되고 있으며, 계산과학은 이론, 실험 등의 보완뿐만 아니라 실험을 대체하거나 새로운 현상 및 이론들을 예측하기 위해 활용되기도 한다. 본질적으로 계산과학 분야에서의 실험적 연구들은 모델링(이론 구축)과 시뮬레이션(구현

및 실행)을 적용하는 형태로 진행된다. 우리는 어떠한 시스템(S)을 잠재적 데이터 소스로 정의하며, S에 대한 모델(M)을 통해 S에서 데이터를 추출하는 과정을 계산과학적 실험으로 정의한다. S에 대한 질문에 답하기 위해 계산 과학자들은 실험(E)의 진행을 위해 다음과 같은 과정들을 수행하게 된다.

실험(E)의 진행:

1. 문제의 인식
2. 문제를 포함하는 시스템(S)의 적절한 개념화(모델(M)의 생성)
3. 시뮬레이션을 위해 S의 연구에 적합한 알고리즘 프레임워크 설계
4. 적합한 컴퓨팅 인프라 선택(일반 병렬컴퓨팅, 그리드 컴퓨팅, 슈퍼컴퓨팅 등)
5. 시뮬레이션 계산 능력의 극대화(하드웨어/소프트웨어적 최적화, 충분한 컴퓨팅 자원)
6. 시뮬레이션 결과와 S 사이의 유사도 평가(M의 검증)
7. 평가 결과에 따른 S의 개념화 조정(M의 수정)
8. 적절한 수준의 M의 검증을 얻을 때까지 6-7 과정을 반복

위와 같은 흐름으로 계산과학 분야에서의 실험적 연구 방법들이 현재 거의 모든 연구 분야들에 걸쳐 활용되고 있으며, 이러한 형태의 접근들을 연구에 활용하는 연구자들 역시 계산 과학자로 간주될 수 있다. 이론적 연구들을 수행하거나 컴퓨터과학(computer science)적 성격의 연구를 수행하는 계산 과학자들을 제외하면, 대부분의 계산 과학자들은 주로 위와 같은 실험적 성격의 계산 과학적 연구들을 수행하고 있다.

2 계산과학 연구 분야

계산과학은 다양한 분야에서 연구를 수행하는 계산 과학자들의 연구 분야들을 포괄하고 있는 과학 분야이고, 계산 과학자는 주로 과학 컴퓨팅(scientific computing, SC)에 능숙한 연구자를 의미한다. 예를 들어, 물리학, 화학 또는 공학 분야 등의 각 응용 분야에서 최첨단 이론, 알고리즘 및 기술들을 개발/발전시키기 위해 다양한 방법들을 고성능 컴퓨팅(high performance computing, HPC)을 통해 적용/연구하는 과학자, 엔지니어, 응용 수학자 등은 계산 과학자로 불릴 수 있다. 최근 계산 과학자들은 거의 모든 과학, 공학 분야들에 걸쳐 시뮬레이션 기반으로 도출된 연구 결과들을 다학제간 활발히 교환하여 타 연구 분야들에서도 새로운 접근법들을 적용/응용할 수 있도록 촉진하고 있다. 계산 과학자들은 정교한 수치 방법, 새로운 계산법 및 데이터 활용법뿐만 아니라 새로운 하드웨어 및 관련 기술들을 기반으로 한 계산 플랫폼들을 개발/활용/공유하면서, 과학, 공학, 산업 등의 영역들에 편재하는 복잡한 다중 규모/영역 문제들의 해결을 목표로 하는, 다학제간 혁신적인 첨단 연구 활동들을 수행하고 있다.

계산과학은 대표적으로 다음과 같은 분야들을 포함한다.

1. 계산과학에 필요한 수학적 방법들의 이론적 개발(예: 수치선형대수(numerical linear algebra), 행렬/텐서 계산(matrix/tensor computation), 반복 계산법(iterative method) 등의 영역에서의 방법 개선 및 새로운 계산법들의 개발)
2. 모델링(modeling), 수치해석(numerical analysis), 알고리즘(algorithms) 및 시뮬레이션(simulations) (예: 수치 및 비수치(numerical and non-numerical), 이산 및 연속(discrete and continuous))
3. 과학(예: 생물학적(biological), 물리학적(physical) 및 사회학적(social)), 공학(engineering), 의학(medicine) 및 인문학 문제(humanities problem) 등을 해결 및 가시화(visualization)하기 위한 소프트웨어(software) 개발
4. 고급 시스템 하드웨어(advanced system hardware), 소프트웨어, 네트워킹(networking) 및 데이터 관리 구성요소들(data management components)을 개발, 활용하고 최적화하는 컴퓨터과학(computer science) 및 정보과학(information science)

위 분야들과 관련된 몇 가지 예들을 나열해 보겠다. 우선 텐서 해석학(tensor analysis) 기반의 희소데이터근사법(data-sparse approximation)이라는 수학적 계산법이 새롭게 개발됨에 따라 계산 과학자들은 기존 방법들을 사용할 때 보다 훨씬 더 적은 계산 비용으로도 수치적으로(numerically) 거의 정확한 주요 고유값(eigenvalue)들을 얻을 수도 있게 되었다. 예를 들어, 희소데이터근사법에 기반한 양자역학계(quantum mechanical system)의 파동함수(wave function) 모델링(modelling)을 통해 고유값 문제를 해결하게 되면 문제에 따라 $O(n^d)$ 의 계산 비용을 요구하는 고유값 문제를 $O(dr^2n)$ ($r \ll n$)의 계산 비용으로도 가장 지배적인(most dominant) 고유벡터(eigenvector)에 대응되는 고유값을 특정 소수 자리까지(예컨대 소수 4~10번째 자리까지도) 수치적으로 정확하게 계산해낼 수도 있다. 따라서, 최근 전산 물리학(computational physics)과 전산 화학(computational chemistry) 분야에서는 수치적으로 거의 정확한, 고유 상태(eigenstate)에서의 가장 낮은 에너지(lowest energy) 및 그 위의 몇몇 들뜬 상태(excited state)들에 대응되는 에너지들을 효율적으로 계산하기 위해 위와 같은 계산법들이 함께 활용되고 있다.

그리고 수치해(numerical solution)의 계산을 위한 계산법 및 알고리즘 - 란초스 알고리즘(Lanczos algorithm)을 통한 에르미트 행렬(Hermitian matrix)의 고유값과 고유벡터(eigenvector)의 계산, 테일러 급수(Taylor's Series)의 수렴(convergence) 응용, 자동 미분(automatic differentiation)을 이용한 미분 계수 계산, 룽게-쿠타 방법(Runge-Kutta method)을 이용한 상미분방정식의 해 계산, 몬테카를로 방법(Monte Carlo method), 이산 푸리에 변환(discrete Fourier transform) 등 - 들을 이론적으로 직접 개발/개선하는 연구자들의 노력은 계산 과학자들이 수행하고자 하는 수치해석(numerical analysis) 작업들 및 이들이 구하고자 하는 수치해(numerical solution)들을 계산하는 작업들을 보다 효과적으로 수행해낼 수 있도록 도와준다.

계산과학 소프트웨어 개발과 관련한 계산과학 분야의 예로는, 전산 물리학 및 전산 화학 분야에서의 제일원리시뮬레이션(First-principles simulation)을 손쉽게 수행할 수 있도록 개발된 시뮬레이션 소프트웨어(simulation software)들이 물리, 화학 분야들을 포함한 여러 연구 분야들에서 널리 사용되고 있다. 제일원리시뮬레이션 소프트웨어들은 계산 과학자들로 하여금 원자, 분자 스케일의 물리, 화학계들을 어렵지 않게 모델링(modeling) 할 수 있도록 개발되어 있을 뿐만 아니라, 여러 물리적 파라미터(parameter)들을 조정해가며 다양한 물리적

조건 아래에서 시뮬레이션을 수행할 수 있도록 해준다.

컴퓨터과학(computer science)적 성격을 띠는 계산과학 분야의 예로는, 대규모 시뮬레이션(large-scale simulation)을 위한 슈퍼컴퓨팅(supercomputing) 인프라 구축 및 클라우드 계산 플랫폼(cloud computing platform)들의 개발/효율화를 위한 연구 분야를 예로 들 수 있고, 이는 다른 계산 과학자들로 하여금 더욱 큰 규모의 시뮬레이션을 수행할 수 있도록 도와준다.

위와 같이 계산 과학자들은 이론적인 영역에서부터 소프트웨어 및 하드웨어적인 영역에 이르기까지 다양한 분야에서 연구를 수행하고 있다. 또한, 한 계산과학 분야에서의 연구의 산물이 다른 계산과학 분야의 발전을 직접적으로 촉진하는 경우들도 상당히 많다.

3 계산과학 방법 및 알고리즘

계산과학에 사용되는 다양한 알고리즘과 수학적 방법들 중 대표적으로 다음과 같은 방법들이 자주 활용된다.

1. 기호적 계산(symbolic computation)을 포함하는 컴퓨터 대수(computer algebra)
2. 유한 차분들(finite differences)에 의한 도함수들(derivatives)의 계산을 포함한 수치 해석(numerical analysis)
3. 수렴 점근 전개(convergent and asymptotic series)로의 테일러 급수(Taylor series) 응용
4. 자동 미분(automatic differentiation, AD)에 의한 도함수 계산
5. 편미분 방정식(partial differential equation, PDE)의 수치적 해를 얻기 위한 유한요소법(finite element method, FEM)
6. 테일러 급수 및 리차더슨 외삽(Richardson extrapolation)을 통한 고차 미분 근사(high order difference approximations)

7. 직사각형 공식(rectangle rule), 사다리꼴 공식(trapezoid rule), 심프슨 공식(Simpson's rule)과 같은 균일한 메시(uniform mesh)에 대한 적분법들(methods of integration)
8. 상미분방정식들(ordinary differential equations, ODE)을 풀기 위한 룬게-쿠타 방법(Runge-Kutta method)
9. 뉴턴법(Newton's method)
10. 이산 푸리에 변환(discrete Fourier transform)
11. 몬테카를로 방법(Monte Carlo(MC) methods)
12. 수치 선형 대수(numerical linear algebra)
13. 선형 프로그래밍(linear programming)
14. 분기 절단법(branch and cut)
15. 분기 한정법(branch and bound)
16. 분자 동역학(molecular dynamics, MD), 칼-파리넬로 분자 동역학(Car-Parrinello molecular dynamics)
17. 공간 매핑(space mapping)
18. 동적 시스템에 대한 시간전진법(time stepping methods for dynamical systems)

그리고 계산과학의 기반이 되는 대표적인 프로그래밍 언어(programming languages), 계산 소프트웨어(computational softwares), 수치라이브러리(numerical libraries), 병렬프로그래밍 라이브러리(libraries for parallel programming), 그리고 현대 계산과학에서 그 역할이 매우 중요하게 강조되고 있는 기계학습/딥러닝을 위한 라이브러리들(libraries for machine learning/deep learning)의 예를 나열해 보면 다음과 같다.

1. 프로그래밍 언어(programming languages)
 - 1) 포트란(FORTRAN(FORMula TRANslator))
 - 2) C, C++

3) 파이썬(Python)

4) R

5) 펄(Perl)

6) 줄리아(Julia)

7) 루아(Lua)

8) 하스켈(Haskell)

2. 계산 소프트웨어(computational softwares)

1) 옥타이브(Octave)

2) 매트랩(MATLAB)

3) 매쓰매티카(Mathematica)

4) 메이플(Maple)

5) 싸이랩(Scilab)

6) TK Solver

3. 수치 라이브러리(numerical libraries)

1) ALGLIB

2) Basic Linear Algebra Subprograms(BLAS)

3) Linear Algebra PACKage(LAPACK)

4) Intel Math Kernal Library(Intel MKL)

4. 병렬 프로그래밍 라이브러리(libraries for parallel programming)

1) Open Multi-Processing(OpenMP)

2) Message Passing Interface(MPI)

- 3) Open Computing Language(OpenCL)
- 4) Compute Unified Device Architecture(CUDA)
5. 머신러닝/딥러닝 라이브러리(libraries for machine learning/deep learning)
 - 1) 텐서플로우(TensorFlow) 및 케라스(Keras)
 - 2) 토치(Torch), 파이토치(PyTorch)
 - 3) 인지 툴킷(Cognitive Toolkit, CNTK)
 - 4) 티아노(Theano)
 - 5) 아파치 MXNet(Apache MXnet)
 - 6) 체이너(Chainer)
 - 7) 카페(Caffe)
 - 8) 사이킷런(Scikit-learn)
 - 9) Azure Machine Learning(Azure ML)
 - 10) Open Neural Network Exchange(ONNX)
 - 11) 엔비디아 텐서RT(NVIDIA TensorRT)

4 계산과학 응용 분야

계산과학은 컴퓨터 프로그램의 개발/활용을 통해 복잡한 수식들을 계산하거나 현상(phenomenon)들을 모사(simulation), 예측(prediction)하는 학문이다. 소립자의 상태와 움직임을 예측하거나 나노 스케일의 분자들로 구성된 물리화학적 시스템을 모사하기도 하며, 보다 거시적으로는 자동차 충돌 시의 차체 왜곡이라던가, 비행기 주변의 기류, 건축 구조물의 안정성, 폭발 장치의 위력, 은하계 별의 움직임 등을 예측하기 위해 컴퓨터 시뮬레이션을 활용하기도 하며, 도시 설계, 날씨 예측 및 기상 이변으로 인한 자연재해를 예측하기도 한다. 그뿐만 아니라, 사회적 현상, 경제적 현상의 분석과 예측 및 의학, 금융, 예술, 체육 연구 분야들에서의 수치해석 및

시뮬레이션 기법들의 활용 사례들도 어렵지 않게 찾아볼 수 있다.

계산과학적 접근을 통한 연구 방법들은 다수의 연구 결과들을 통해 그 효율성이 실증적 사례들을 통해 검증되었으며, 1980년대부터 과학기술 연구 방법의 한 축으로 자리 잡게 되었다. 1990년대 보잉사는 보잉 777의 개발을 위해 보잉 747과 보잉 767의 정보를 기반으로 한 컴퓨터 시뮬레이션을 통해, 보잉 777의 개발 기간을 대폭 단축할 수 있었고, 국내 자동차, 건설 회사들도 컴퓨터를 통한 설계와 시뮬레이션의 도움을 받아 큰 프로젝트들을 성공적으로 완수하고 있다. 과학의 영역에서도 계산과학의 도움이 없이는 현재의 나노과학 및 유전학이 달성한 학문적 수준에 도달할 수 없었을 것으로 학계에서는 판단하고 있다. 아인슈타인의 일반상대성 이론은 중력파가 존재한다고 예측하고 있는데, 1990년대 전산 물리학적 방법을 통해 중력파의 존재성에 대한 이론을 뒷받침할 수 있었으며, 최근 이론 및 계산과학을 통해 예측되었던 중력파가 실제로 관측됨에 따라 이에 큰 기여를 하였던 물리학자들에게 2017년 노벨물리학상이 수여되었다. 또한, 계산과학 소프트웨어를 개발한 과학자들도 노벨상을 수여하거나 노벨상 후보에 자주 오르고 있다.

계산과학 시뮬레이션 프로그램들은 실제적 현상 속에서 실제로 변화하는 조건들을 모델링하고 있고, 많은 경우 ‘논리적 메시(logical mesh)’들을 직접 생성, 활용하면서 시뮬레이션을 수행하는 형태들을 취하고 있다. 예를 들어, 기상 모델에서 다루는 공간을 나누고 있는 각 메시는 평방킬로미터(square kilometer) 스케일일 수 있을 텐데, 이들 각 공간 영역을 표현하는 메시는 이들 메시들의 수만큼 분할된 컴퓨터 메모리의 각 영역에 매핑(mapping), 각 메모리 분할 영역들은 실제 공간적 메시가 가지고 있는 위치, 고도, 풍향, 습도, 온도, 압력 등의 정보들을 지니도록 하였다고 생각해 보자. 이때, 시뮬레이션 프로그램은 현재 상태를 기반으로 약간의 시간이 흐른 뒤의 가능한 다음 상태를 계산하는 방식으로 작동하도록 할 수 있을 것이며, 이러한 계산은 미분 방정식들의 수치해(numerical solution)를 구하는 방식으로 주로 진행된다. 수많은 연구 분야들 속에서 이와 같은 흐름으로 현상을 예측하거나 이해를 시도할 수 있는 문제들은 셀 수없이 많을 것이다. 대표적으로, 다음과 같은 계산과학의 응용 사례들이 자주 소개가 되고 있다.

1. 예측 계산과학(predictive computational science)

예측 계산과학은 초기 및 경계 조건(initial and boundary condition)과 일련의 특성화 매개 변수(parameter) 및 관련 불확실성(uncertainty)을 고려하여 물리적 이벤트의 특정 측면들을 예측하도록 설계된 수학적 모델들을 생성/수정하거나 수치해를 계산하거나 검증하는 등의 연구들을 수행하는 과학 분야이다.

2. 도시 복잡계(Urban complex systems)

도시는 엄청난 복잡성으로 인해 모델링, 분석, 이해 및 관리에 더 많은 노력이 필요한 대규모의 시스템이다. 도시는 사람들의 사교성에서부터 문화적 구조, 교통 시스템, 도시 계획에 이르기까지 수많은 다양한 요소들이 강력하게 얽혀있는 시스템으로 볼 수 있다. 도시의 각 요소 하나하나도 이미 그 자체로도 복잡한 시스템이긴 하여, 이들이 상호 얽혀있는 도시 복잡계 문제는 인간이 다루고 있는 가장 복잡한 문제들 중 하나로 간주되고 있다. 특히, 전 세계 인구의 절반 이상이 도시에 살고 있음과 앞으로 도시로 인구가 몰리는 추세는 더욱 강해질 것이라는 예측과 함께, 미래의 도시 개발을 예측, 이해, 구체화하기 위해서, 가능한 재난 상황들에 대한 대비를 위해서 효과적인 계산 모델과 시뮬레이션 기법들의 발전이 필요하다는 목소리가 높아지고 있다. 도시 복잡계 연구의 핵심은 복잡계적 방법 등을 포함한 다양한 모델링과 시뮬레이션들을 통해 도시 동역학(city dynamics)에 대한 이해를 높임과 함께, 앞으로 다가올 더욱 큰 규모의 도시화도 함께 대비하는 것이다.

3. 전산 금융(computational finance)

오늘날의 금융 시장에서 막대한 양의 상호 의존적 자산은 서로 다른 위치와 시간대에 있는 수많은 상호 작용하는 시장 참여자들에 의해 거래되고 있다. 이들의 행동 양상들은 전례가 없을 정도의 복잡성을 가지고 있고, 금융 시장에서 발생하는 현상들은 일반적으로 매우 복잡한 형태의 수학적 계산 모델들을 통해 표현된다. 이러한 모델들은 풀어내는 것은 보통 효율적인 수치 알고리즘들의 도움 없이는 해결이 불가능하다. 특히, 시간이 흐름에 따라 유동성과 신용 등으로 인해 일어날 수 있는 상호 의존적 위험 요소들에 대한 고려 또한 더욱 중요해지고 있기에, 다중스케일(multi-scale)의 전체론적 접근(holistic approach) 방식을 통한 현상의 이해가 요구되고 있는 상황이며, 이러한 복잡계를 수학적으로 최대한 효과적으로 표현한 후, 막대한 컴퓨팅 자원을 활용하는 계산과학적 접근을 통해 복잡한 시스템 대한 이해 및 현상 예측들을 시도하게 된다.

4. 전산 물리학/화학(computational physics/chemistry)

전산 물리학은 과학 분야에서 현대 컴퓨터를 최초로 적용하였던 분야이다. 물리학에서 수학적 모델에 기반한 다양한 이론들은 시스템이 어떻게 작동하는지에 대한 매우 정확한 예측을

제공하지만, 모델 자체를 푸는 것이 불가능할 정도로 어려운 경우가 많다. 전산 물리학은 수치적 해석 및 근사를 통해 복잡한 시스템들을 이해하기 위한 분야로, 이론 물리학의 하위 분야 또는 이론 물리학과 실험 물리학의 중간 지점의 분야로 간주되기도 한다. 전산 물리학은 이론과 실험을 모두 보완하는 연구 영역이다. 전산 물리학이 다루는 범위는 굉장히 광범위하다. 소립자 스케일에서의 물리적 현상들을 컴퓨터 시뮬레이션을 통해 다루는 격자 양자 색역학(lattice quantum chromodynamics, lattice QCD)에서부터 응집물리학(condensed matter physics), 생물물리학(biophysics), 유체역학(fluid mechanics), 천체물리학(astrophysics)에 이르기까지 전산 물리학의 범위는 광범위하다. 또한, 물리학적 이론들에 기반한 전산 화학(computational chemistry)이 다루는 연구 영역들은 전산 응집물리학/고체물리학 등에서 다루는 연구 영역들과 겹치는 경우가 많다.

전산 화학(computational chemistry)은 컴퓨터 시뮬레이션을 통한 계산을 통해 이론화학의 문제들을 다루는 화학의 한 분야이다. 이론화학 문제들은 일반적으로 분자, 원자, 전자들로 구성된 화학적 시스템을 물리학적 방정식들(슈뢰딩거 방정식, 뉴턴 운동 방정식, 통계 물리학적 방정식 등)을 통해 모델링 되어 표현된다. 이렇게 표현된 모델을 계산과학적 방식으로 그 해를 구하여 연구 대상으로 삼았던 화학적 시스템을 이해한다. 전산 화학 또한 괄목할만한 컴퓨터 성능의 향상으로 인해, 이론과 실험 분야와 어깨를 나란히 할 만큼의 확고한 제3의 연구 분야로 자리 잡게 되었다. 전산 화학은 주로 다음과 같은 방법들을 통해 화학 문제들을 다룬다.

- 1) 분자궤도함수 이론(molecular orbital, MO)
- 2) 분자동역학(molecular dynamics, MD)
- 3) 몬테카를로 방법(Monte Carlo(MC) method)
- 4) 분자역학(molecular mechanics, MM)
- 5) 밀도범함수 이론(density functional theory, DFT)

5. 전산 생물학(computational biology)

최근 생명 과학 및 공학 분야에서의 일어나고 있는 놀라운 연구 성과들은 의학 연구 분야에까지 큰 혁명을 일으키고 있다. 예컨대, 대량 시퀀싱(high-throughput sequencing), 실시간 중합효소 연쇄반응(real-time polymerase chain reaction, RT-PCR), 세포 내

이미징(intra-cellular imaging), 유전자 발현의 현장 부합법(in-situ hybridization), 가벼운 시트 형광 현미경(light sheet fluorescence microscopy, LSFM) 및 광학적 투사(optical projection)와 같은 3차원 이미징 기술(three-dimensional imaging techniques), 컴퓨터 단층 촬영(computer tomography) 등에 의해 생성되고 있는 방대한 양의 복잡한 데이터들이 축적되고 있는 상태에서, 최근 인공지능(artificial intelligence, AI) 기술의 괄목할만한 성장과 함께 컴퓨터가 새로운 단백질 구조들을 인간을 뛰어넘을 정도의 놀라운 수준으로 예측해내는 등 빅데이터(big data) 기반의 전산 생물학적 분석/예측 결과들이 의약학 연구들에도 직접적인 이바지를 하고 있다. 대부분의 생물학적 시스템들은 복잡계로 인식되어 주로 다중스케일 모델(multi-scale model)의 개발을 통해 표현되며, 전산 생물학적 문제들을 수치적으로 해결하기 위해 막대한 수준의 컴퓨팅 자원을 필요로 하는 경우들이 많다.

6. 복잡계 이론(complex systems theory)

일반적인 수치 모델링적 접근을 통한 컴퓨터 시뮬레이션으로 해결할 수 없는 수준의 계산 복잡도를 지니는 문제들을 복잡계적 접근 방식을 통해 시스템을 이론적으로 모델링하여 다루는 연구 분야이다. 복잡계에 대한 해석은 각 학문 분야와 학자들마다 다소 상이하나, 복잡계에 대한 공통적인 해석은 다음과 같다.

“현실은 너무나도 복잡하여 방정식이나 단순한 논리 체계로는 환원시키는 게 불가능하다.”

복잡계의 이론적 모델링의 첫 번째 축은 분석의 미시적/거시적 관점이다. 현상을 분석할 때 미시적 움직임/행태에 집중하는 경우와 거시적인 변화(각 변수들로 인해 시스템 전체가 어떻게 변화하는지)에 집중하는 경우로 나뉜다.

복잡계의 이론적 모델링의 두 번째 축은 분석방식이다. 분석방식은 모든 변수/메커니즘을 동일하게 일반화시켜 계산하는 방식과 각각의 변수/메커니즘에 차이를 두어 계산하는 경우로 나뉜다. 예를 들어, 복잡계 네트워크를 통한 복잡 행위자 모델, 게임이론 등을 통한 합리적 선택 모델, 비선형 시계열 분석을 통한 해석적 매크로 모델은 복잡계 모델이다.

7. 계산과학 및 공학(computational science in engineering, CSE)

계산과학 및 공학(CSE)은 공학적 분석 및 설계에서 발생하는 복잡한 물리적 문제들을 해결하기 위한 모델 및 시뮬레이션 알고리즘을 개발/활용하는 분야이다. CSE는 컴퓨터 시뮬레이션을 통해

기존의 실험적 방법으로는 수행할 수 없던 실험을 대신하거나 기존의 실험들을 통해 얻었던 정보들을 시뮬레이션에 활용하는 등의 접근 방식들을 통해 산업 현장에서 경제적, 시간적 비용을 줄이는데 크게 기여하고 있다.

참고문헌

A. M. Turing (1937), 「On Computable Numbers, with an Application to the Entscheidungsproblem」, Proceedings of the London Mathematical Society, Volume s2-42, Issue 1, Pages 230-265.

John Vincent Atanasoff (1940), 「Computing Machine for the Solution of Large Systems of Linear Algebraic Equations」, https://digitalcollections.lib.iastate.edu/islandora/object/isu%3AAtanasoff_142

A. M. Turing (1950), 「Computing machinery and intelligence」, Mind, Volume LIX, Issue 236, Pages 433-460.

J. Dongarra 외(2000), 「Guest Editors Introduction to the top 10 algorithms」, Computing in Science & Engineering, Volume 2, Pages 22-23.

Haohuan Fu 외 (2016), 「The Sunway TaihuLight supercomputer: system and applications」, Sci. China Inf. Sci., Volume 59, Pages 072001.

Azalia Mirhoseini 외 (2021), 「A graph placement methodology for fast chip design」, Nature, Volume 594, Pages 207-212.

J. Jumper 외(2021), 「Highly accurate protein structure prediction with AlphaFold」, Nature, Volume 596, Pages 583-589.

K. Tunyasuvunakool 외(2021), 「Highly accurate protein structure prediction for the human proteome」, Nature, Volume 596, Pages 590-596.

T. Schlick 외(2021), 「Biomolecular modeling thrives in the age of technology」, Nat. Comput. Sci., Volume 1, Pages 321-331.

SamuelFosso Wamba 외(2021), 「Are we preparing for a good AI society? A bibliometric review and research agenda」, Technological Forecasting and Social Change, Volume 164, Pages 120482.

Barry A. Cipra(2000), 「The Best of the 20th Century: Editors Name Top 10 Algorithms」, SIAM NEWS.

관계부처합동(2021), 「4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략」, 비상경제 중앙대책본부.

미래와 과학(2017), 「‘고양이 얼굴’ 알아본 인공지능, 움직임까지 예측」, 한겨레 신문
서진근 외, 4차 산업혁명과 산업수학, 2017_06호, NRF Issue paper

양정모, 산업수학 전략과제 종료 이후의 새로운 융합연구사업 제안, 2020 대한수학회 가을 발표회 proceeding

양정모, 2021년 메시지(더 나은 세상을 위한 수학)가 주는 시사점, NRF R&D 브리프, 2021-5호

이상남 외, 2020년도 국가과학기술표준분류체계의 개선 및 활용기반 구축, 2020-027, KISTEP

조승한 기자(2021), 「구글 딥마인드AI, 인간 단백질 2만개 구조 모두 예측했다」, 동아사이언스, 2021.07.23.

Hyperion Research (2020)

TOP500, 「TOP500 LIST - JUNE 2021」.

위키백과, 「Computational science」.

위키백과, 「슈퍼컴퓨터」.

위키백과, 「AlphaFold」.

나무위키, 「복잡계 연구방법론」.

저자소개

| 최 성 수

물리학 박사, 전산물리학/계산과학전공

| 양 정 모

수학 박사, 한국연구재단 정책혁신팀장

| 이 준 업

감수, 수학 박사, 한국연구재단 자연과학단장

본 이슈리포트의 내용은 NRF의 공식적인 의견이 아닌 집필진의 견해이며 동 내용을 인용 시 출처를 밝혀야 합니다.

NRF ISSUE REPORT 2021_18호

병렬컴퓨팅 발전에 따른 계산과학 분야들의 발전과 확장,
그리고 연구 패러다임 변화

| 발행일 | 2021년 10월 12일

| 발행인 | 이 광 복

| 발행처 | 한국연구재단

본 원 : 대전광역시 유성구 가정로 201

서울청사 : 서울특별시 서초구 현릉로 25

<http://www.nrf.re.kr>

| 편 집 | 정책연구실 정책혁신팀

ISSN 2586-1131