



LLM Sandwich: A Neuro-Symbolic Approach to Solving Complex Reasoning Problems

Jennifer Chu-Carroll



Preaching to the Choir: LLMs Cannot Reason!

Faithful Chain-of-Thought Reasoning.

Retrieval-Augmented Reasoning for Commonsense Reasoners.

Least-to-most prompting enables complex reasoning in large Language Models.

Large Language Models are reasoners with Self-Verification.

LAMBADA: Backward Chaining for Automated Reasoning in Natural Language.

Language Models are Multilingual Chain-of-Thought Reasoners.

Language Models are Zero-Shot Reasoners.

... A Unified Approach.

" ...the main thing that they [LLMs] are bad at is reasoning. And a lot of the valuable human things require some degree of **complex reasoning**."

- Sam Altman, OpenAI CEO

Hard Fork podcast, November 20, 2023

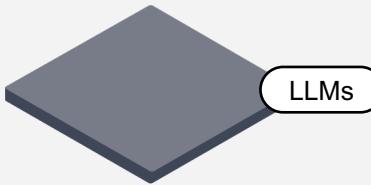
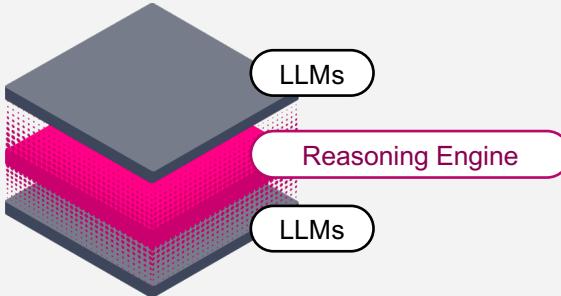


Reasoning + LLMs = ?



Reliable complex reasoning based on **formal and efficient mathematical algorithms**.

- Mathematical constraint modeling
- Efficient constraint propagation and backtracking
- Constraint identification, resolution, and repair
- Possible worlds reasoning and optimization
- Dependency management & truth maintenance
- Constraint-space decomposition and layering
- Linear and non-linear programming
- Operations research
- Bayesian analysis



Low stakes use cases

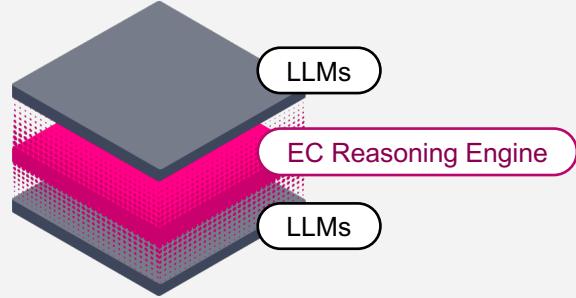
Generate, transform and summarize content using statistical models.

Limitations of LLMs:

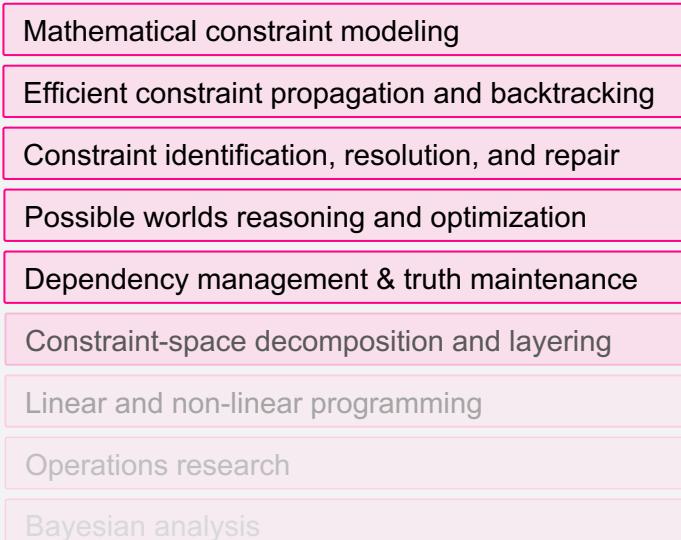
- Hallucinate answers
- Fail at general complex reasoning
- Not grounded
- Lack transparency
- High training costs



LLM Sandwich: A Neuro-Symbolic Approach to Solving Complex Reasoning Problems



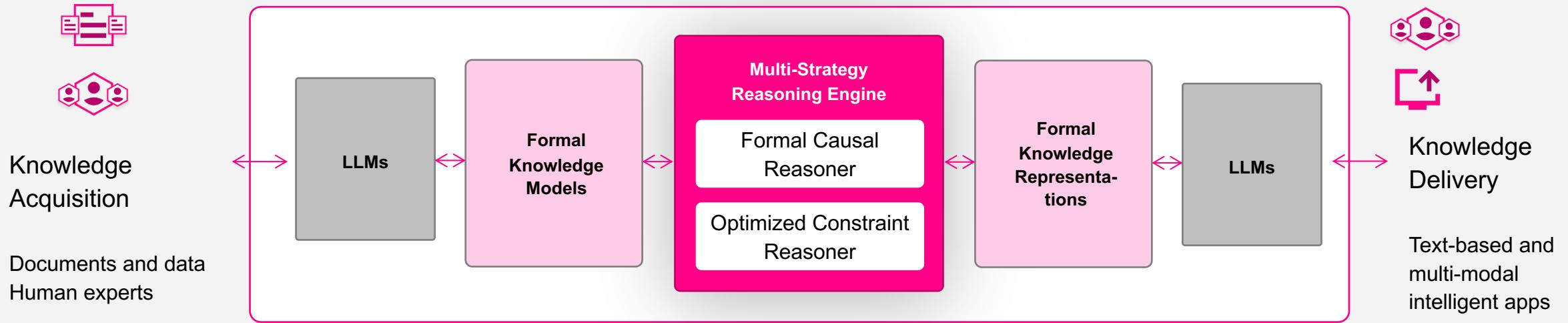
Interactive, natural-language-enabled complex reasoning capability



- Interactively and reliably solve practical and difficult problems
- Get provably-correct answers across domains
- Enable total decision transparency
- Eliminate need for highly-skilled developers while improving reliability
- Accelerate time to market for robust AI applications
- Scale stable and efficient AI implementations
- Much lower run-time compute costs



A Closer Look at the LLM Sandwich



↔ **Fluent**
Interact with AI in natural language.

✓ **Reliable**
Get predictable, valid, and unbiased results.

☰ **Transparent**
See explicit, traceable decision procedures.

TargetException **Precise**
Use unambiguous formal systems.

⌚ **Efficient**
Compute efficiently with formal algorithms.



Cogent: Holding the Sandwich Together

Formal Language as Natural Language

↔ No programming required

Precisely capture business knowledge in unambiguous English (*Cogent*) so anyone can read and understand it

↔ Executable specifications

Abstract specifications are directly processed by EC's general-purpose reasoning engine

↔ Single source of truth

Changes made to English specifications automatically update the app logic, re-generate APIs and deploy a new version

The screenshot shows a mobile application interface titled "Staff Scheduling". At the top, there is a navigation bar with a logo and three horizontal dots. Below the title, there are two sections: "# Definitions" and "# Business Rules".

Definitions

We say that a shift is "**covered by**" a nurse "**on**" a day if the nurse is working the shift on that day.

We say that a nurse is "**booked on**" a day if the nurse is working a shift on that day.

Business Rules

Each shift **must** be covered by exactly one nurse on each day.
A nurse **cannot** be working more than one shift on the same day.
Each nurse **must** be booked on at least two days.
A nurse **may** be on vacation on a day.
A nurse **cannot** be working a shift on a day if the nurse is on vacation on that day.

The screenshot shows a mobile application interface titled "Architectural Optimization". At the top, there is a navigation bar with a logo and three horizontal dots. Below the title, there are two sections: "# Definitions" and "# Business Rules".

Definitions

A "**cooking center**" is a type of work center.
A "**cleanup-prep center**" is a type of work center.
A "**refrigeration center**" is a type of work center.

Business Rules

When a cooking center combines a refrigeration center with a cleanup-prep center, this is called a "**work triangle**".
A cooking center **may** "combine" a refrigeration center "**with**" a cleanup-prep center.
Every refrigeration center **is required to be** against a wall.



Knowledge-Authoring Assistance

Capture precise business knowledge as executable natural language in *Cogent* using generative AI

↔ Capture knowledge with precision

Interact with assistant to capture application knowledge as a precise and formal specification

↔ Edit with writing assistant

Get real-time guidance on how to conform with precise and unambiguous language

↔ Generate test data

Accelerate validation with auto-generated test data for a given *Cogent* specification

Model Builder

How can i add a business rule stating that there may be different departments nurses may work in?

First, define '**Department**' as a new type.

Then, associate a department with each shift so that the model knows which department is relevant for a given shift.

Here's how you can update the model to include departments:

Department is a type.
Departments are "**Pediatrics**", "**Emergency**", "**Oncology**", and "**Surgery**".
A shift may be "**in**" a department.
A nurse may work in a department.

Nurse Scheduling

Day **is** a type.
Shift **is** a type.
Nurse **is** a type.
A nurse **may** be working a shift on a day.
The shifts **are** morning, evening, and night.
"Monday", **"Tuesday"**, **"Wednesday"**, **"Thursday"** and **"Friday"** are days.

Department **is** a type.
Departments **are** "**Pediatrics**", "**Emergency**", "**Oncology**", and "**Surgery**".
A shift **may** be "**in**" a department.
A nurse **may** work in a department.



Conversational Knowledge Acquisition

App Preview

↔ Predefined Templates for Verticals

Generic meta-models for select verticals that capture the common rules/constraints in those application areas

↔ Guided Dialogue to Specialize

Dialogue assistant prompts developer for application specific information to instantiate their model

↔ Zero Shot Dialogue Interaction

Zero shot user dialogue allows developer to interactively verify system behavior in real time

The screenshot shows a conversational interface titled "Model Builder Assistant". The developer's messages are in blue boxes, and the AI's responses are in grey boxes.

- Developer:** Hello! I'm here to help you configure your staff scheduling application. Why don't you start by telling me a little about your business?
- AI:** I am responsible for medical staff scheduling at North Lake Trauma. We have emergency physicians to diagnose and treat critically injured patients as well as general surgeons for emergency surgery. We are staffed with both surgical and ER nurses.
- Developer:** I understand that you're in charge of medical staff scheduling at North Lake Trauma. Your team consists of emergency physicians, general surgeons, surgical nurses, and ER nurses. I've updated the technical specification document to reflect these roles. Could you please tell me more about the shifts at your workplace? For example, how many shifts are there in a day and what are their timings?
- AI:** 4 Line Changes
- Developer:** We have 4 shifts a day: 6am-noon, noon-6pm, 6pm-midnight, and midnight-6am.
- AI:** I understand that your workplace operates in 4 shifts: 6am-noon, noon-6pm, 6pm-midnight, and midnight-6am. I've updated the technical specification document to include these shifts. Next, I'd like to understand the qualifications required for each position. Could you please tell me if there are any specific qualifications or certifications required for the emergency physicians, general surgeons, surgical nurses, and ER nurses?
- Developer:** Qualifications are taken care of elsewhere. Let's talk about staffing. The number of ER physicians and nurses we require is based on the average patient intake volume. This means we typically need at least one physician |

Staff Planner ...

The user may choose for an employee to be unavailable on a day.

When an employee works as a position during a shift on a day, the cost of this is 2.

Rules of Inference and Constraints

> Working as a position entails being scheduled for a day

An employee is scheduled during a shift on a day if the employee is working as a position during the shift on the day.

> To work as a position the employee must be qualified.

An employee cannot be working as a position during a shift on a day unless the employee is qualified for that position.

> Enforce user preferences

An employee must be working as a position during a shift on a day if the employee is requested to work as the position during the shift on the day.

An employee must be scheduled during a shift on a day if the employee is requested to be scheduled during the shift on the day.

An employee cannot be scheduled during a shift on a day if the employee is requested unscheduled during the shift on the day.

Editable Constraints

> These are the default constraints for a staff scheduling problem.

> Ask the author if each of these constraints makes sense for their domain.

At least one employee must be working as each position during each shift on each day.

An employee cannot be scheduled during more than one shift on the same day.

Custom Application

The days are Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday. "emergency physician" is a position.

"general surgeon" is a position.

"surgical nurse" is a position.

"ER nurse" is a position.

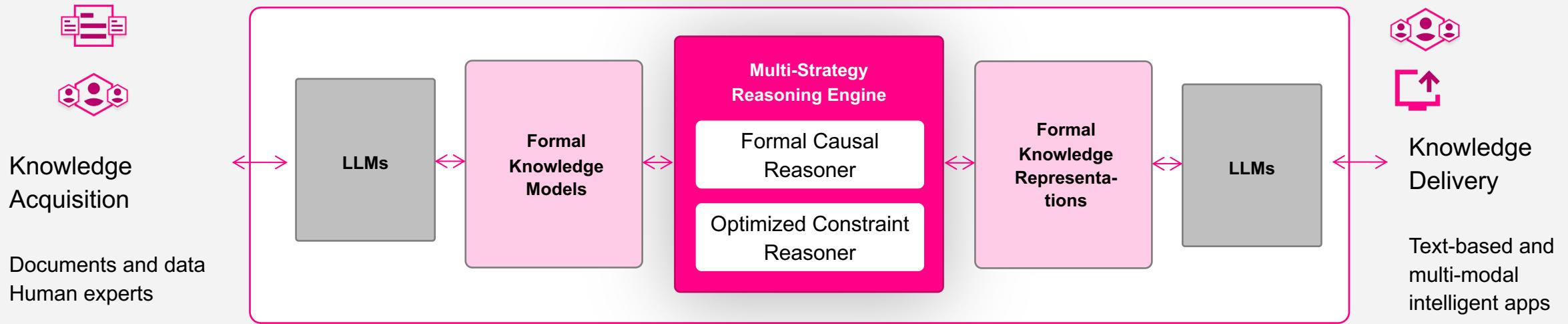
"6am-noon" is a shift.

"Noon-6pm" is a shift.

"6pm-midnight" is a shift.

"Midnight-6am" is a shift.

Knowledge Acquisition vs. User Interaction



 **Fluent**
Interact with AI in natural language.

 **Reliable**
Get predictable, valid, and unbiased results.

 **Transparent**
See explicit, traceable decision procedures.

 **Precise**
Use unambiguous formal systems.

 **Efficient**
Compute efficiently with formal algorithms.



Interactive application generation

Generate multi-modal app interfaces that combine the fluency of LLMs with the rigor of formal reasoning.

↔ Auto-generate APIs

APIs to interact with the reasoning engine automatically generated and deployed. Integrate with data and build multi-modal apps from deployed APIs

Staff Scheduler

Your desired plan is not possible because Andy is on vacation.

The cheapest way to optimize your schedule and guarantee it works is to swap shifts for Chris and Andy. Does this work?

View Schedule

That works. Can I get an alternative plan with four suggestions on call too?

Project Planner

Can I assign Adam to P4, he is the best to kick this project off?

P4 is set to start on Week 2 and Andy is already assigned to P1 until the end of Week2.

But we can find a solution. Here are your options:

Assign John to P4 on Week 2

Start P4 on Week 2

Plan

	Week 1	Week 2	Week 3
P1	Andy, Chris		
P2		Don, Betty	
P3	Don		Adam
P4		John	Chris

Roster

	Morning	Evening	Night
Mon	Andy	Chris	Betty
Tue	Chris	Betty	Andy
Wed	Don	Chris	
Thu	Chris	Betty	
Fri	Don	Chris	

Live Chat

Configured an optimal plan that meets all your requirements and is ready to book. Is there anything else you would like to change?

I need to be in London for June 15

Ok, got you in London for June 15. I also found better pricing with fewer layovers on a March 26 departure, want to explore the options?

Sounds great! 😊

Nadi, Fiji

Tokyo, Japan

London, United Kingdom

12 Jun → 20 Jul 8D

Lima, Peru

New York, USA

24 Aug

↔ Auto-generate conversational apps

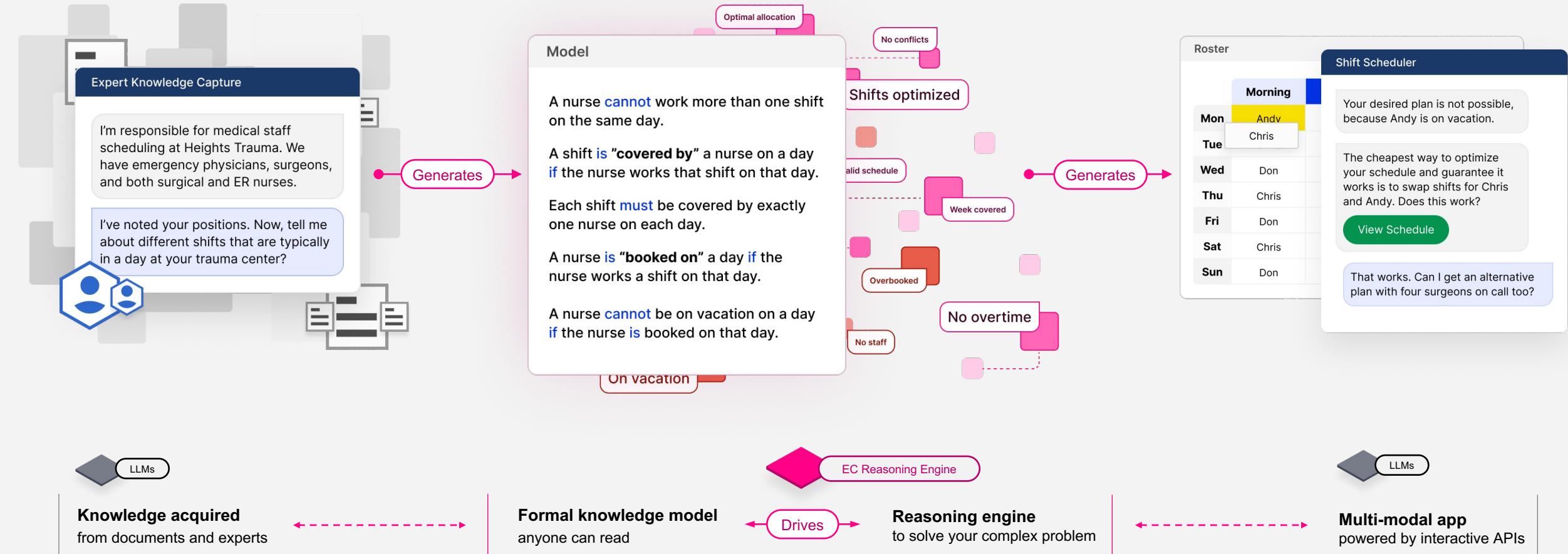
Generate conversational applications out of the box with EC Reasoner-to-LLM interfaces.

↔ Easily evolve and update over time

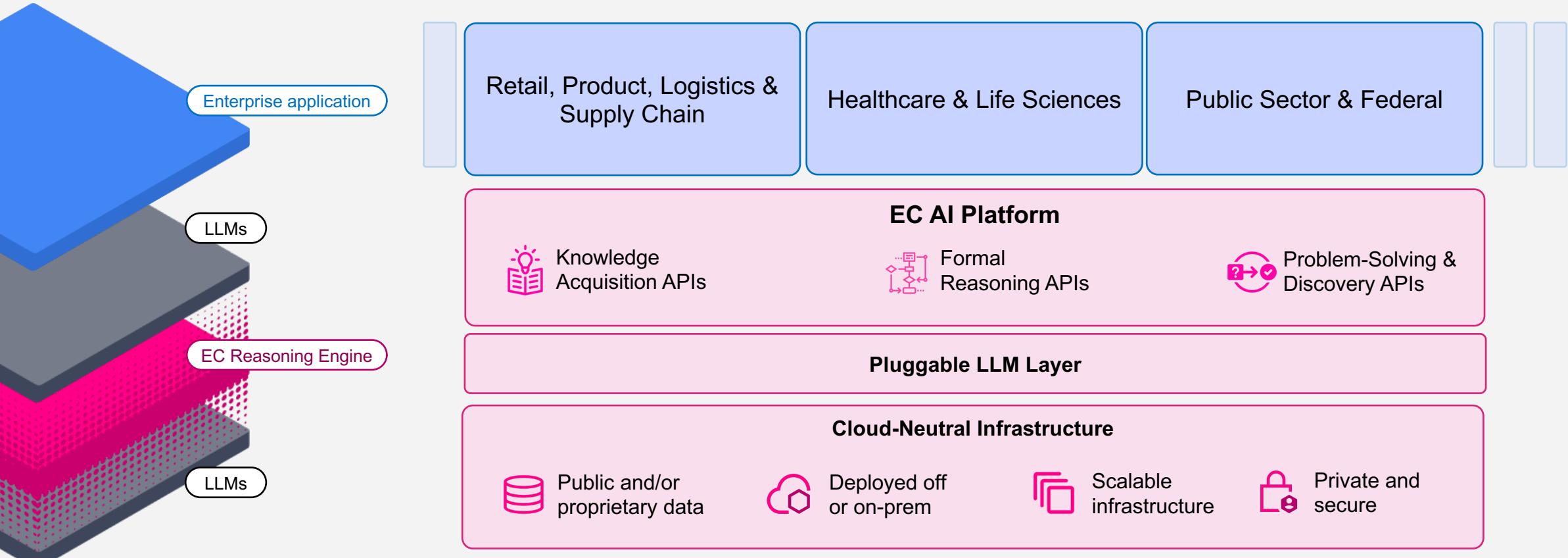
Automatically update APIs and conversational UIs whenever you make updates to business rules in Cogent.



Putting The Sandwich Together



Scalable Enterprise AI Platform for Reliable, Complex Reasoning



Application in Complex Travel Planning

- Complex fare and travel rules and restrictions
- Constrained pricing model
- Soft and hard customer requirements
- Real-time flight availability across multiple airlines
- Changing requirements must be negotiated with user
- Fast, reliable re-planning to accommodate problems
- Huge search space, few satisfy all constraints at any given time



Customer Preferences

Put Istanbul between Athens and Rome
I want to spend 2 weeks in Tokyo and need to be in Lima on May 28th

Travel may not be via the point of origin.

Only one intercontinental departure and one intercontinental arrival permitted in each continent except as follows:

1. Two permitted in North America.
2. Two permitted in Asia when one is for travel between the Southwest Pacific and Europe/Middle East.
3. Two permitted in Europe/Middle East for travel to/from/via Africa. If travel is to/from Europe in both directions, itinerary may

Only one international departure and one in permitted.

EXCEPTION: Two permitted for origin USA without stopover.

NOTE: travel between USA and Canada is not counted as international.

Resource Availability

Any two cities may be connected by flights on any given date... or not

Changes in real time



Production Example of Interactive Optimization in Travel and Logistics



Your itinerary

Base price: \$5,699 USD (+ taxes and fees)

Travelling for 60 days

New York, US	15 May	
Mexico City, MX	15 May - 23 May	8 days
Sydney, AU	25 May - 2 Jun	8 days
Nadi, FJ	2 Jun - 10 Jun	8 days
Tokyo, JP	11 Jun - 19 Jun	8 days
London, GB	19 Jun - 27 Jun	8 days
Lima, PE	27 Jun - 5 Jul	8 days
Miami, US	5 Jul - 13 Jul	8 days

mapbox

The world map displays a travel route starting in North America (New York), moving to South America (Mexico City, Lima), then to Oceania (Sydney, Nadi), Asia (Tokyo), Europe (London), and finally back to North America (Lima, Miami). The route is indicated by a dashed line connecting the cities.

Alaska AIRLINES American AIRLINES BRITISH AIRWAYS CATHAY PACIFIC FINNAIR IBERIA JAPAN AIRLINES malaysian QANTAS QATAR AIRWAYS

Type a message...

POWERED BY ELEMENTAL COGNITION GIVE FEEDBACK

How long would you like your entire trip to span?

2 months

Ok, I have set your trip duration to 60 days

Here are my suggested departure and arrival dates.

Would you like to make any changes?

4:38 PM

I want to spend 2 weeks in Tokyo and need to be in Lima on April 1st

Ok, I have noted you need to be in Lima on April 1st, 2024

Ok, you have requested to stay in Tokyo for 14 days

Would you like to make any more date changes?

I'm ready to lookup flights

Optimizing date assignments...

Production Example of Interactive Optimization in Degree Planning

The screenshot displays a web-based degree planning application interface. At the top, there's a navigation bar with tabs for 'Default path' and 'New Path +'. Below this, a toolbar includes 'Course' (selected), 'Bachelor of Business', 'Major', 'Accounting', 'Minor', 'No minors', 'Edit', 'Undo', '24/144 Requirements', and a search bar for 'finance'. A sidebar on the right shows a list of courses with filters applied for 'Clayton' and '6 CP'. The main area is divided into sections for different academic years and semesters, each containing a grid of course cards. Courses are represented by boxes with titles, descriptions, credit points (e.g., 'A | 6 CP'), and status indicators (checkmarks or 'Swap' buttons). Some courses have 'Search for unit' buttons. A blue callout box provides a tip: 'Tip: you can click on a unit to see more information.' Another box contains instructions for free electives: 'For your remaining 7 free electives ...' followed by a bulleted list: '• A maximum of 3 may be Level 1', '• A minimum of 1 must be Level 3', and '• The remaining 3 must be Level 2 or 3'. It also notes that up to 60 credit points can be taken outside the Caulfield campus.

Default path : New Path +

Course Bachelor of Business Major Accounting Minor No minors Edit Undo 24/144 Requirements

Open in student portal

finance

Filters

17 results Only show compatible

S1-01 Clayton 6 CP

show more...

ECB3121 Add

Economics of international trade and finance

This unit develops an understanding of the theory of international trade. It explains how international trade creates gains for a nation and the effect this has on its economic outcomes, such as...

S1-01 Peninsula 6 CP

show more...

ECC3810 Add

Public finance

The relationship between the public and private sectors is explored from the perspective of both public finance and political economy. Theories of public expenditure and revenue generation are...

S1-01 Clayton 6 CP

show more...

Remove from path

ECF2721 Remove

Trade finance and foreign exchange

This unit focuses on the topics of international finance and trade. This includes the importance of international finance, the determination of exchange rates, foreign investment, parity...

S1-01 Caulfield 6 CP

show more...

ACF2400 Swap 2

Integrative thinking in a dynamic world

A | 6 CP

ACF2100 Swap 2

Financial accounting

B | 6 CP

ECF2721 Swap 2

Trade finance and foreign exchange

C | 6 CP

COVERED A FREE ELECTIVE

Search for unit

Tip: you can click on a unit to see more information.

For your remaining 7 free electives ...

- A maximum of 3 may be Level 1
- A minimum of 1 must be Level 3
- The remaining 3 must be Level 2 or 3

If desired, you may take up to 60 credit points outside of the Caulfield campus

ACF2400 Swap 2

Management accounting

B | 6 CP

Free elective

Search for unit

ACF3200 Swap 2

Management accounting

B | 6 CP

Free elective

Search for unit

ACF3100 Swap 2

Advanced financial accounting

B | 6 CP

ACX3300 Swap 2

Predictive analytics in accounting

B | 6 CP

Free elective

Search for unit

Free elective

Search for unit

ETC3520 Add

Foundations of quantitative finance

This unit will introduce you to the basic theories of financial market behaviour, modelling of investment risk, interest rate models and option pricing models. You will be introduced to the mathematical...

S2-01 Clayton 6 CP

show more...

BEX2421 Add

Foundations of quantitative finance

This unit will introduce you to the basic theories of financial market behaviour, modelling of investment risk, interest rate models and option pricing models. You will be introduced to the mathematical...

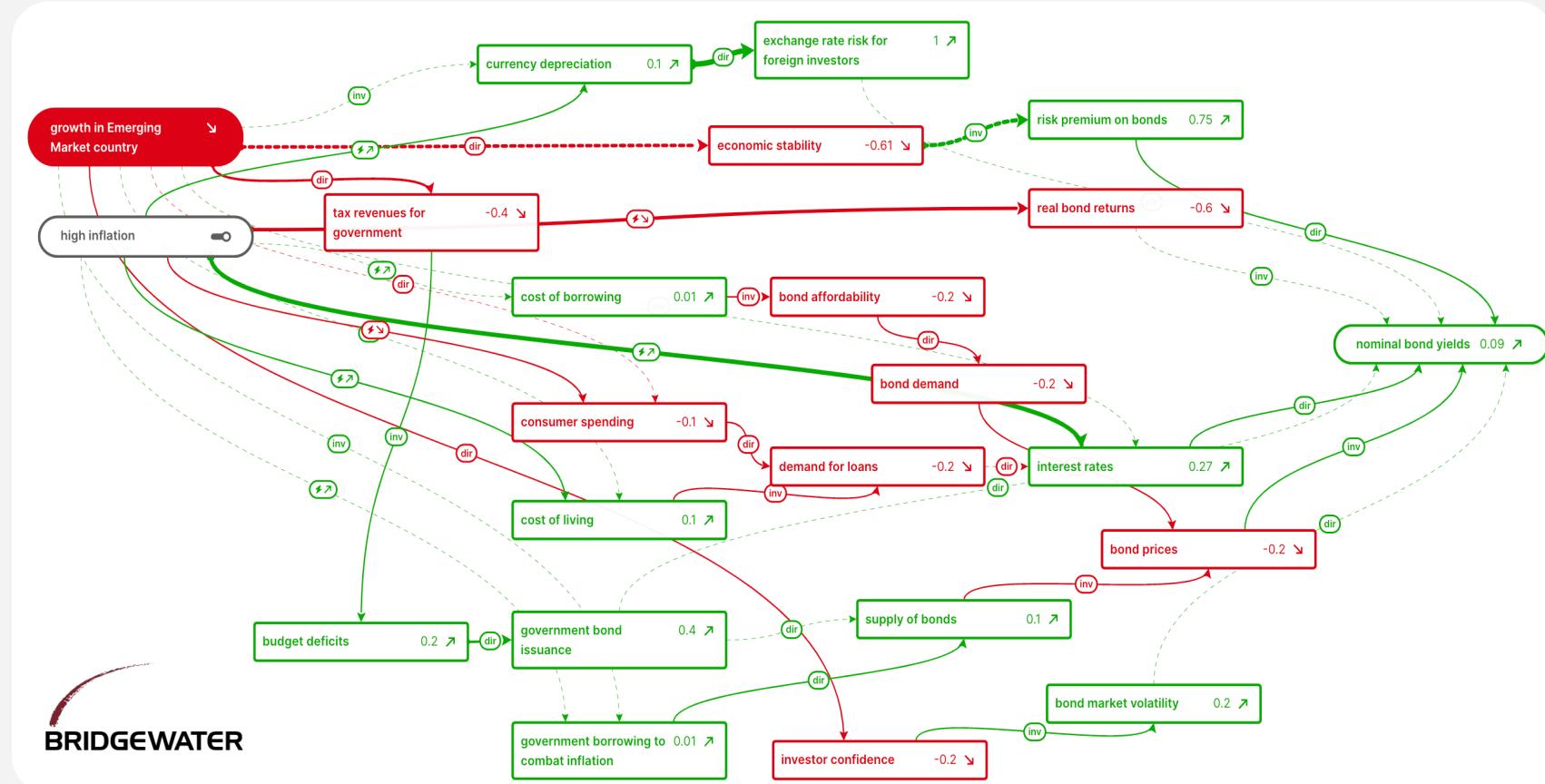
S2-01 Clayton 6 CP

show more...

Production Example of Causal Reasoning in Macroeconomics

The economic growth in an Emerging Market is slowing down. The country is facing high inflation. **Should we bet on nominal bond yields rising?**

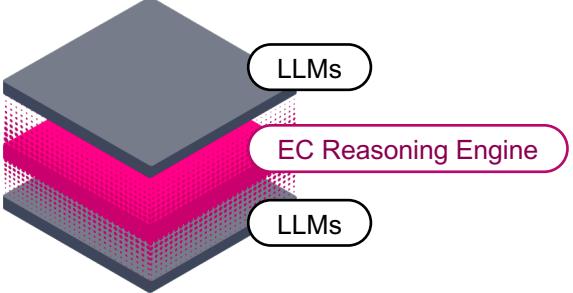
- Generate robust causal models
- Consistently propagate causal influence with mathematical fidelity
- Perform complex, rigorous and traceable What-if Analysis
- Deliver higher quality research and better justified decisions even on small data



EC AI performs **formal logical and causal analysis** of alternative hypotheses given the evidence, before it generates an answer.



Beating a Dead Horse: LLMs Cannot Reason!

-  : combines the reliability of a general-purpose formal reasoning engine and the flexibility of language processing of LLMs
- Deployed systems demonstrated the capabilities of applications built on the EC AI platform
- Our evaluation shows EC systems to significantly outperform GPT-4 on complex travel planning, workforce planning, and degree planning

Report: <https://ec.ai/performance-benchmarks/>