From Generating Answers to Building Explanations

Integrating Multi-Round RAG and Causal Modeling for Scientific QA

Victor Barres, Clifton McFate, Aditya Kalyanpur, Kailash Kartik Saravanakumar, Lori Moon, Nati Seifu, Abraham Bautista-Castillo

Corresponding Authors: Victor Barres (victor.barres@gmail.com) ; Clifton McFate (mcfateclifton@gmail.com)

What Characteristics Make Research Questions Difficult to Answer?

- Use Case: research queries in life sciences:
  - *How does epigenetic dysregulation of neurotrophins impact AD risk?*
  - *How is the loss of RHO linked to podocyte function?*
- Characteristics of Research Questions
  - Causal and Predictive
  - Knowledge Intensive
  - Multi-document
  - Multi-step inference
  - Multiple correct answers
  - Open Domain

Are LLMs Sufficient?

- LLMs End to End
  - LLMs are opaque and prone to hallucination
  - Retraining takes time and resources; they rapidly become out of date in fast moving fields
- Retrieval Augmented Generation
  - Retrieve documents from an external trusted source and prompt the LLM using these documents
  - Reduces hallucinations and makes it easy to update with new documents
  - Reasoning is needed to guide search and integrate findings
- Chain of Thought Prompting
  - The LLM generates intermediate "reasoning" steps to keep it on track and provide traceability
  - Logical validity isn't guaranteed

Beyond an Answer: What is Missing?

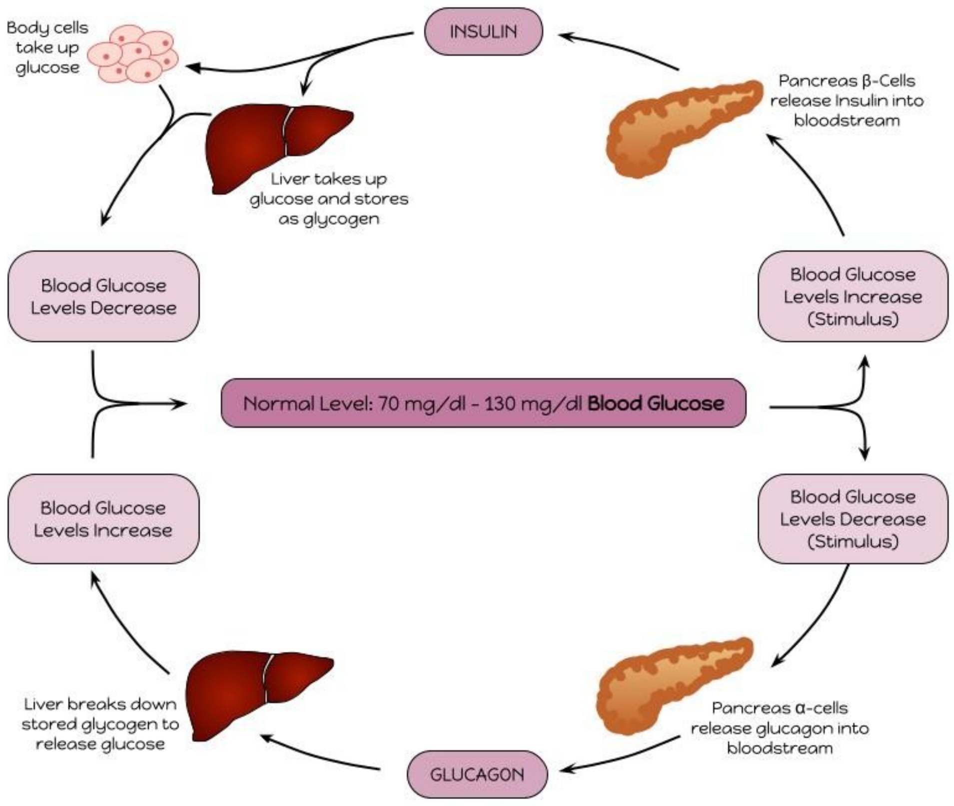
- Miller (2019) summarizes a long history of social sciences research, suggesting four criteria of explanatory answers.
  - Selective
    - There may be infinite ways to explain a phenomenon at arbitrary levels of granularity
    - Good explanations are based around a consistent, often restrictive, causal lens
  - Mechanistic rather than probabilistic
    - A factor simply being associated with increased or reduced likelihood of an event is rarely satisfactory for human explanations
  - Contrastive
    - Good explanations show why an outcome occurs as opposed to a specific alternative
  - Transactional
    - Explanations are social, and the causal lens is negotiated between interlocutors

How Can RAG Better Fit These Criteria?

- Select a causal lens to guide and constrain RAG
  - Should reflect a causal lenses commonly used in research
  - Should be mechanistic, not associative
- Guide RAG by iteratively generating edges in a causal graph from retrieved documents
  - Directed graphs are an intuitive interface for causal models, facilitating user editing and enabling graph based reasoning
  - Iterative generation allows precise user feedback during explanation generation

Designing a Causal Lens

- Qualitative Process Theory captures how people qualitatively model and reason about dynamic processes
  - **Processes** (*insulin release*) govern change in **quantities** (*insulin*) via **direct influences** which propagate via **indirect influences** to other quantities (*blood glucose level*)
- Our causal graph adopts QP-like influence edges and adds discrete states
  - Nodes:
    - **States**: Boolean statements about the world
    - **Quantities:** Any fluent that changes over time
  - Edges:
    - **Influence:** A quantity influences another quantity, positively or negatively
    - **Triggers**: States can trigger or be triggered by state and quantity changes

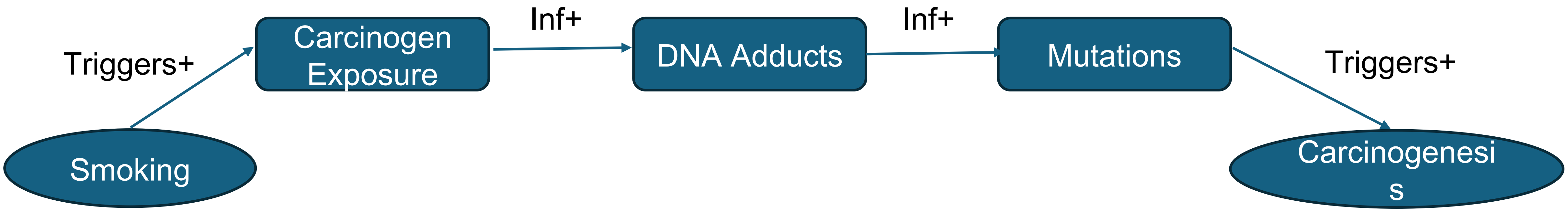


Example Illustration of Glucose Homeostasis

Image Attr: Wikipedia - https://commons.wikimedia.org/wiki/File:Glucose_Homeostasis_for_Wikipedia.jpg

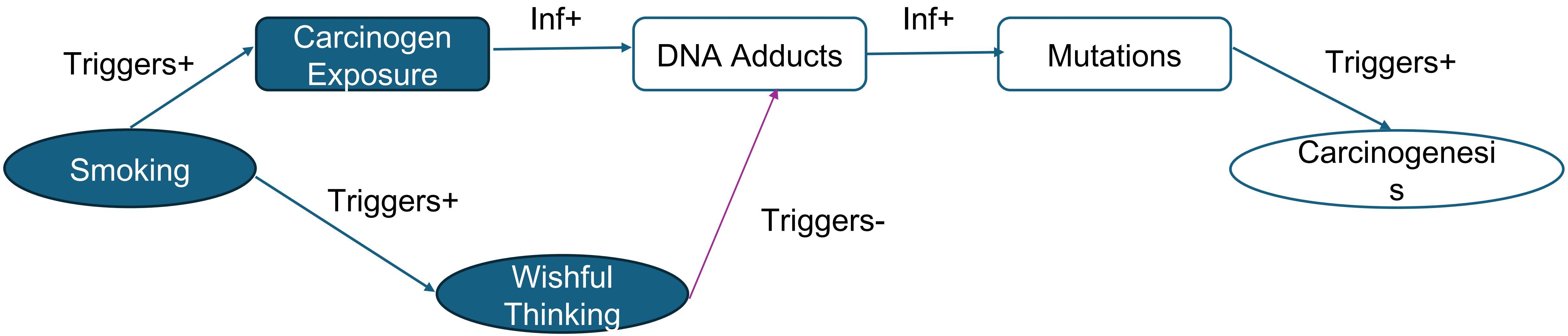
What Reasoning does the Solution Graph Support?

- Expressed as constraints for an Answer Set Programming model implemented via our Cogent reasoning engine
- **Quantities** can be {-1, 0, 1} : decreasing, unknown, increasing
- **States** can be {-1, 0, 1}: false, unknown, true
- Each node's sign is determined by the sum of its influences



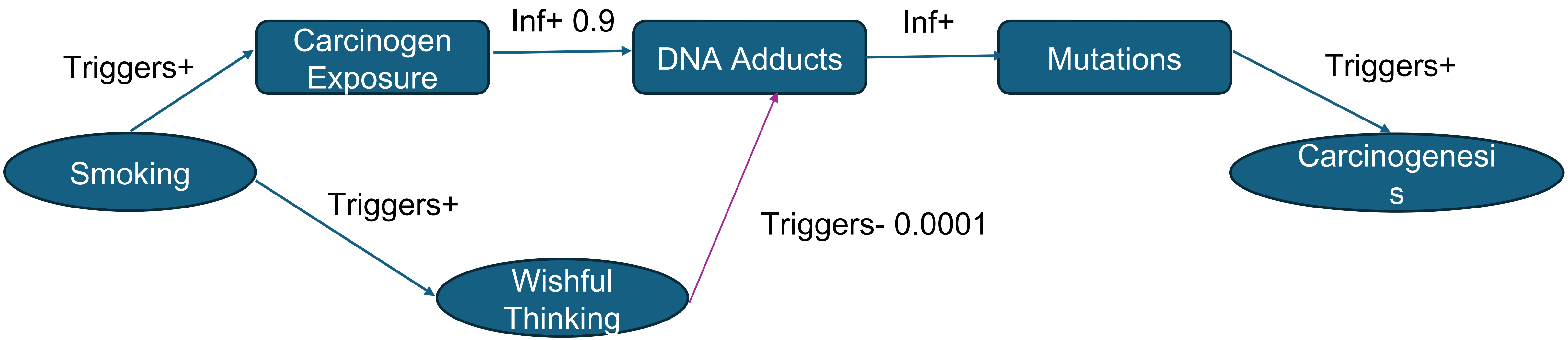
What Reasoning does the Solution Graph Support?

- Expressed as constraints for an Answer Set Programming model implemented via our Cogent reasoning engine
- **Quantities** can be {-1, 0, 1} : decreasing, unknown, increasing
- **States** can be {-1, 0, 1}: false, unknown, true
- Each node's sign is determined by the sum of its influences



What Reasoning does the Solution Graph Support?

- Expressed as constraints for an Answer Set Programming model implemented via our Cogent reasoning engine
- **Quantities** can be {-1, 0, 1} : decreasing, unknown, increasing
- **States** can be {-1, 0, 1}: false, unknown, true
- Each node's sign is determined by the sum of its influences



QP Backed Solution Graph

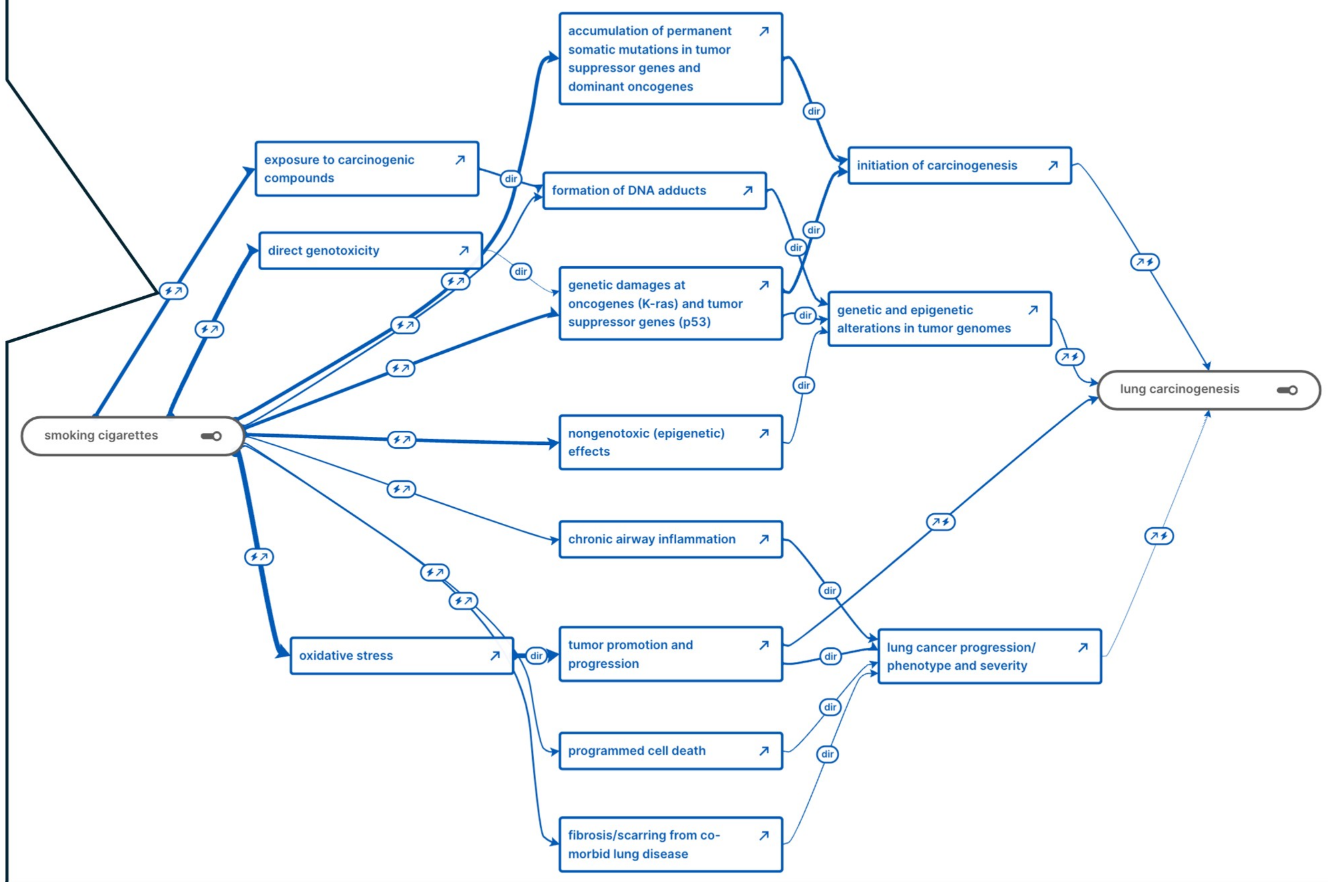
Does smoking cigarettes **trigger increase of** exposure to carcinogenic compounds?

Evidence (34)

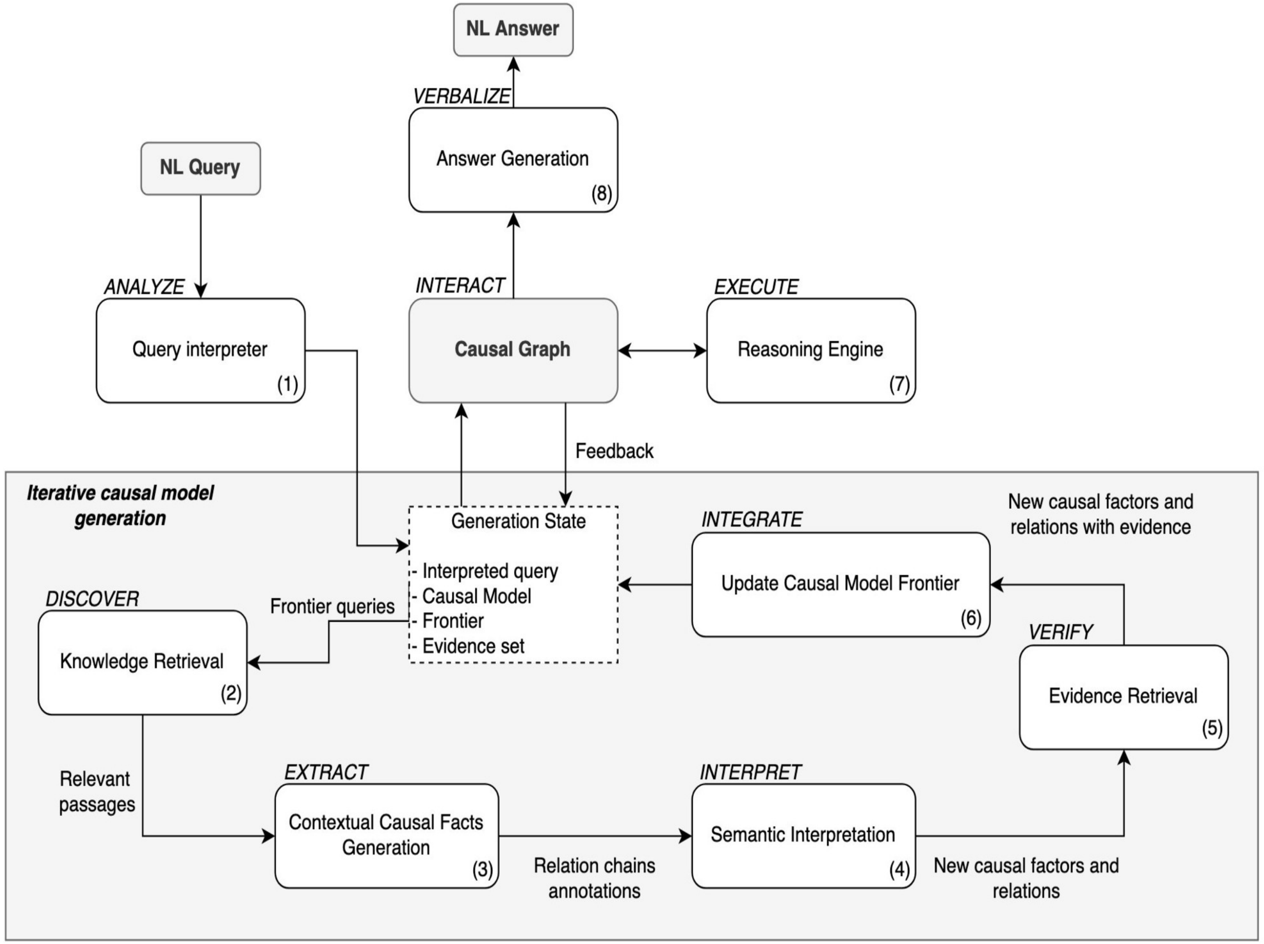
There is ample evidence that the risk of developing various types of cancer is greater among smokers than non-smokers (Grando, although the role of smoking in the etiology of some cancers remains controversial (Li et al.; Shao et al.. Tobacco smoke contains at least 4,000 compounds (nicotine and a number of carcinogenic compounds such as tobacco-specific nitrosamines, polycyclic aromatic hydrocarbons, and aldehydes) that are capable of inducing the DNA damaging response that initiates tumorigenesis and enhances the spread of metastases. The risk and severity of cancer depend on the duration of exposure and the amount of tobacco smoking (Schane et al..

Dozens of toxic chemicals are either added to cigarettes or produced during tobacco combustion. Many of these compounds are carcinogenic, and some chemicals are added to increase product appeal, ease smoking initiation, discourage cessation, or promote relapse.

Cigarette smoke (CS) has been established as one of the major risk factors for many pathologies including lung cancer in humans and experimental animals. In view of the discrepancy about the role of alpha-tocopherol (AT) in carcinogenesis, the present study was designed to investigate the effects of different doses of AT on benzo(a)pyrene-DNA [B(a)P-DNA] adduct formation in lungs of CS inhaling mice. Extent of



Iterative Graph Building



- Query Interpreter extracts independent and dependent factors, becoming frontiers of the graph
- At each turn
  - Retrieve documents for causal links between frontiers
  - Extract QP causal annotations with evidence
  - Expand graph
    - Allow user to edit graph if desired
- Generate NL solution given graph, evidence, and RE labeling

Example: QP Causal Annotation Format

TEXT	**ANNOTATION** **---** **SOLUTION GRAPH RELATION**
Under stress, the body experiences elevated cortisol levels which increase blood pressure	[change=increase] cortisol levels ==CAUSE==> [change=increase] blood pressure --- INFLUENCE(cortisol levels, blood pressure)
Physical activity/exercise interventions have been proven to reduce cellular oxidative stress	[change=increase] exercise ==CAUSE==> [change=decrease] cellular oxidative stress --- INV-INFLUENCE(exercise, cellular oxidative stress)
TIMP-2-defficient mice exhibit increased monocyte/macrophage infiltration	TIMP-2-defficiency ==CAUSE==> [change=increase] macrophage infiltration --- TRIGGERS+(TIMP-2-defficiency, macrophage infiltration)

Evaluation

- Results from 25 end user queries
- Multi-phase answer evaluation: Correctness
  - **Claim Density**: Average information in answer
  - **Citation Density:** Average # verification options / claim
  - **Source Hallucination:** % of hallucinated citations
  - **Citation Rate:** % of claims with at least one valid citation
  - **Justification Rate:** % of cited claims that are justified by source
  - **Relevance Rate:** % of cited, justified claims, that are relevant to the question
- Complexity
  - **Max Number of Hops:** Max Depth
  - **Number of Concepts:** Breadth

Results

System	Claim Density	Citation Density	Source Hallucination Rate	Citation Rate	Justification Rate	RelevanceRate	Max Hops	Avg # Concepts
GPT4-Turbo[1]	4.16	1.01	31.4%	64.42%	27.88%	22.12%	**2.5 ± 2.1**	5.1 ± 3.1
Perplexity[2]	4.76	0.59	0.01%	32.77%	17.65%	11.76%	1.5 ± 1.2	4.0 ± 3.3
Elicit[3]	5.00	**1.36**	0.01%	98.40%	86.40%	60.80%	0.8 ± 0.6	3.3 ± 3.2
Ours	**5.36**	1.14	**0.00%**	**98.51%**	**90.30%**	**86.57%**	2.1 ± 0.7	**7.5 ± 2.4**

1. openai.com
2. https://www.perplexity.ai
3. https://elicit.com/

Conclusions and Future Work

- Humans generate and evaluate explanations with causal lenses
  - Lenses are restrictive and choosing the correct lens is a social process
- We can use a causal lens to guide RAG
  - Specifically, a digraph provides guidance as well as a user interface
- QP Theory facilitates high-level causal explanations
  - However, QP Theory isn't the only causal lens available
  - Ross (2021) proposes a "pathway" model as an alternative to mechanistic causation
  - Users should be able to customize and experiment with different underlying causal models

Citations

- Kenneth D Forbus. 1984. Qualitative process theory. Artificial intelligence, 24(1-3):85–168.
- Kenneth D Forbus. 2019. Qualitative representations: How people reason and learn about the continuous world. MIT Press
- Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267:1–38.
- Lauren N. Ross. 2021. Causal Concepts in Biology: How Pathways Differ from Mechanisms and Why It Matters. The British Journal for the Philosophy of Science, 72(1):131–158. Publisher: The University of Chicago Press.