MONASH University

Formal Explainability for Artificial Intelligence in Dynamic Environments

Jaime Cuartas Granada

26-February-2026

Department of Data Science and Artificial Intelligence (DSAI), Monash University, Australia

Project Summary & Refinements

Summary

Thesis Title: Formal Explainability for Artificial Intelligence in Dynamic Environments.

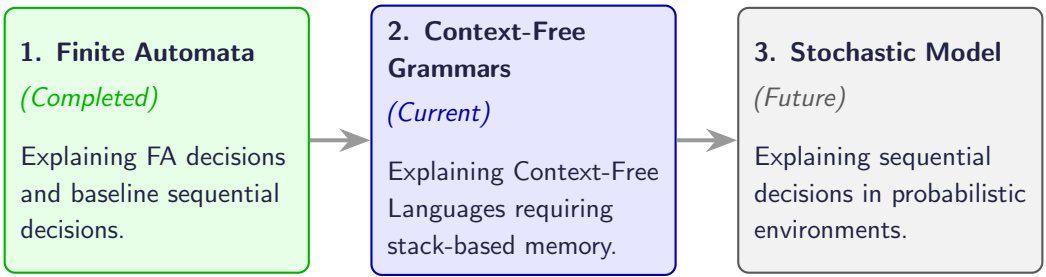
Goal: Deliver explanations for sequential decision-making models.

The Research Roadmap:

1. **Finite Automata (Completed):** Explaining Finite Automata decisions and baseline sequential decisions.
2. **Context-Free Grammars (Current):** Explaining languages Context-Free Languages requiring (stack-based memory).
3. **Stochastic Models (Future):** Explaining sequential decisions in probabilistic environments.

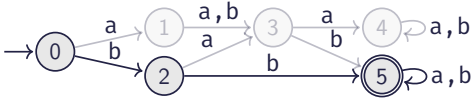
Summary

Goal: Deliver explanations for sequential decision-making models.



Completed Work: Finite Automata

Explaining Finite Automata



Input w**:**	b b b b b	$\rightarrow$ Accept
AXp 1:	b b Σ Σ Σ	$\rightarrow$ Guarantees Accept $L(\mathtt{bb}\Sigma\Sigma\Sigma) \subseteq L(\mathcal{A})$
AXp 2:	Σ Σ b Σ Σ	$\rightarrow$ Guarantees Accept $L(\Sigma\Sigma\mathtt{b}\Sigma\Sigma) \subseteq L(\mathcal{A})$
CXp:	Σ b Σ b b	$\rightarrow$ Enables Reject $L(\Sigma\mathtt{b}\Sigma\mathtt{bb}) \not\subseteq L(\mathcal{A})$

Outcomes: Enumeration & Milestone

- **Algorithmic Contribution:**
  - Leveraged this duality to develop algorithms for the formal **enumeration** of explanations in Finite Automata.
  - Successfully maps the abstract concepts of XAI onto rigorous formal language properties.

Phase 1 Milestone Achieved

Status: Completed.

Output: The formal definitions, duality proofs, and enumeration algorithms have been compiled and submitted to **ICALP 2026**.

- *Transitioning to Phase 2:* With the baseline for regular languages established, we now scale the complexity to languages requiring memory.

Project Refinements Since Confirmation

Refining the Scope and Methodology:

- **Shift from Automata Execution to Grammar Membership:**
  - *Initial Concept:* Explaining the execution trace and stack states of Pushdown Automata (PDA).
  - *Refinement:* Framed the problem purely around **CFG membership** ($w \in L(G)$). This allows us to leverage efficient grammar-based parsing algorithms (like CYK) rather than tracking infinite PDA states.
- **Targeted Probabilistic Scope:**
  - *Initial Concept:* Broad application to generic stochastic models.
  - *Refinement:* Narrowed the final phase to focus specifically on Probabilistic Context-Free Grammars (PCFGs) and Markov Models, ensuring the mathematical foundation built in Phase 2 scales directly into Phase 3.

The Minimal Hitting Set Duality

- **The Core Relationship:** Abductive and contrastive explanations share a formal duality [?].
- Every AXp is a minimal hitting set of the complete set of CXps, and vice versa.
- To flip a prediction (CXp), you must free at least one token from every reason that guarantees the current prediction (AXp).

Motivation

- ADD CONTENT

Separated content, e.g., variable explanations

Frame without a title

- Frames do not need to have a title...

Math Expressions

Integrals and Other Expressions

$$\iint_{\partial\Omega} f(x)\mathrm{d}x \in \mathbb{C} \tag{1}$$

$$E = mc^2 \tag{2}$$

$$F = ma \tag{3}$$

m Mass

c Speed of light

Theorem

The following statement is correct

$$\frac{\partial f(\vec{x})}{\partial x_i} = \sum_{l=1}^{L} \cos\left(l\frac{2\pi}{L} + 0\right) \tag{4}$$

Elements

Typography

```
The theme provides sensible defaults to
^^I^^I\emph{emphasize} text, \alert{accent} parts
^^I^^Ior show \textbf{bold} results.
```

becomes

The theme provides sensible defaults to *emphasize* text, **accent** parts or show **bold** results.

Font feature test

- Regular
- *Italic*
- Small Caps
- **Bold**
- ***Bold Italic***
- **Bold Small Caps**
- `Monospace`
- *`Monospace Italic`*
- **`Monospace Bold`**
- ***`Monospace Bold Italic`***

Lists

Items

- Milk
- Eggs
- Potatoes

Enumerations

1. First,
2. Second and
3. Last.

Descriptions

PowerPoint Meeh.

Beamer Yeeeha.

Tables

Table 1: Largest cities in the world (source: Wikipedia)

City	Population
Mexico City	20,116,842
Shanghai	19,210,000
Peking	15,796,450
Istanbul	14,160,467

Blocks

Three different block environments are pre-defined and may be styled with an optional background color.

Default

Block content.

Alert

Block content.

Example

Block content.

Default

Block content.

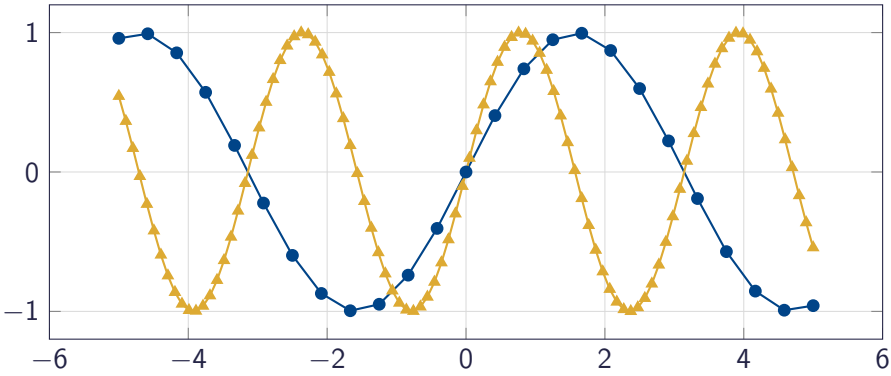
Alert

Block content.

Example

Block content.

Line plots



Standout Frame!

Backup slides

Sometimes, it is useful to add slides at the end of your presentation to refer to during audience questions.

The best way to do this is to include the `appendixnumberbeamer` package in your preamble and call `\appendix` before your backup slides.

The theme will automatically turn off slide numbering and progress bars for slides in the appendix.