Algebraic Machine Reasoning: How inductive reasoning reduces to algebraic computations

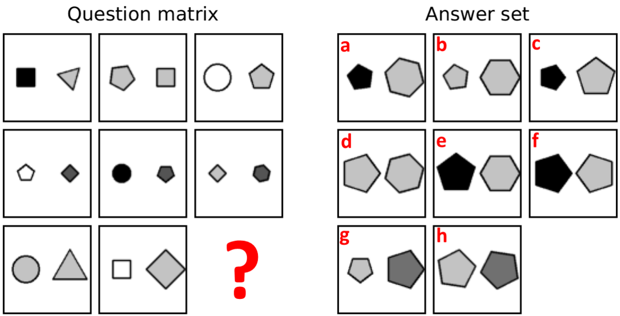
8th Conference on Artificial Intelligence and Theorem Proving
(AITP 2023)

Speaker: Kai Fong Ernest Chong
(Singapore University of Technology and Design)

Based on paper published in CVPR 2023.

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Raven's Progressive Matrices (RPMs)



RPMs are well-known visual-based IQ tests, commonly used to measure *abstract reasoning* and *problem-solving* abilities in humans.

- Question matrix (left): 3×3 grid of panels.
  - First 8 panels are filled with geometric entities.
  - 9-th panel is "missing".
- Answer set (right): 8 panels representing 8 possible answer options.
- **Task:** Determine the *correct* answer for the missing 9th panel.
  - The 9 panels should satisfy some abstract patterns/relations.

Crystallized Intelligence vs Fluid Intelligence

Crystallized Intelligence

- Intelligence derived from **experience and training**.
- Ability to solve problems **previously seen before**.
- Ability to use knowledge in **long-term memory**.

Fluid Intelligence

- Intelligence derived from **mental agility and adaptability**.
- Ability to solve novel problems **not seen before**.
- Ability to infer new relations using **short-term memory**.

What we associate with:

- Stored/learned knowledge (accummulated over time)
- Ontologies, databases
- Requires **lots** of data and prior knowledge
  - Large statistical models
  - Knowledge graphs
  - Large language models

What we associate with:

- Abstract reasoning
  - Logical deductions
  - Discovering new relations (e.g. $E = mc^2$ by Einstein)
- Problem-solving strategies
- Requires **little** prior knowledge
  - Raven's progressive matrices
  - Logic puzzles

Algebraic Machine Reasoning (algMR)

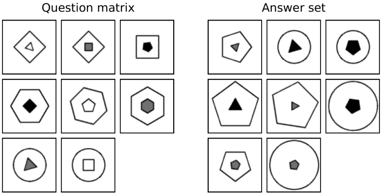
A new reasoning framework well-suited for abstract reasoning.

difficult process of **novel problem-solving** → reduce to → ***routine* algebraic computations.**

- Our framework solves RPMs **without** needing to optimize performance on task-specific data for reasoning.
  - Analogy: A gifted child solves RPMs without needing practice on RPMs.
- What's "algebraic" in algebraic machine reasoning?
  - Commutative algebra: e.g. ideals of rings.
  - Computational algebra: e.g. ideal-based algorithms.
- Key algebraic ideas in algMR:
  - *Define* concepts as **ideals** of a polynomial ring.
  - Inductive reasoning via **primary decompositions** of ideals.
- We solve RPMs effectively as computational problems in algebra!

Reasoning and solving RPMs → becomes → Computing the primary decompositions of ideals

Surpassing human-level performance on RPMs



	Method	Conference	Accuracy
1	ResNet+DRT	CVPR 2019	40.4%
2	LEN	NeurIPS 2019	41.4%
3	Wild ResNet	ICML 2018	44.3%
4	CoPINet	NeurIPS 2019	46.1%
5	DCNet	ICLR 2021	47.0%
6	SRAN	AAAI 2021	60.8%
7	PrAE	CVPR 2021	77.0%
8	**algMR (ours)**	**CVPR 2023**	**93.2%**
	Human	-	84.4%

A typical RPM instance of the I-RAVEN dataset

(See our paper for more baselines and a detailed breakdown of accuracies by RPM configurations.)

Experiment Results

- I-RAVEN dataset: 14,000 RPMs in the test set.
  - RPMs are generated according to 7 configurations.
- Significant accuracy outperformance (directly from raw RPM images).
  - +16.2% **improvement** over previous state-of-the-art accuracy.
  - **Surpasses human-level performance (+8.8% higher)**.

What is an ideal?

- Polynomial ring $R = \mathbb{R}[x_1, \ldots, x_n]$.
  - Set of all polynomials in variables $x_1, \ldots, x_n$ with real coefficients.
  - This set, together with addition and multiplication, forms a **ring**.
  - A monomial is a polynomial with a single term with coefficient 1.
- A subset $I \subseteq R$ is called an ideal of R if there are polynomials $g_1, \ldots, g_k$ in R such that

$$I = \{f_1 g_1 + \cdots + f_k g_k | f_1, \ldots, f_k \in R\}$$

  contains all polynomial combinations of $g_1, \ldots, g_k$.
  - We say that $\mathcal{G} = \{g_1, \ldots, g_k\}$ is a generating set for I.
  - **Notation:** $I = \langle g_1, \ldots, g_k \rangle$ or $I = \langle \mathcal{G} \rangle$.
  - Note: Every ideal has infinitely many possible generating sets.
  - Basic operations on ideals: sums, products, intersections.
- A monomial ideal is an ideal that has a generating set that consists only of monomials.

Definition: A concept of R is a monomial ideal of R.

Examples of concepts

Suppose $R = \mathbb{R}[x_{\text{red}}, x_{\text{blue}}, x_{\text{circle}}, x_{\text{square}}]$.

- Primitive concepts: Concepts generated by a single variable.
  - $\langle x_{\text{red}} \rangle$ is the concept "red".
  - $\langle x_{\text{blue}} \rangle$ is the concept "blue".
  - $\langle x_{\text{circle}} \rangle$ is the concept "circle".
  - $\langle x_{\text{square}} \rangle$ is the concept "square".
- **Note:** $\langle x_{\text{red}} \rangle$ contains several concepts:
  - $\langle x_{\text{red}} \rangle \supseteq \langle x_{\text{red}} x_{\text{circle}} \rangle$ ("red circle").
  - $\langle x_{\text{red}} \rangle \supseteq \langle x_{\text{red}} x_{\text{square}} \rangle$ ("red square").
  - $\langle x_{\text{red}} \rangle \supseteq \langle x_{\text{red}} x_{\text{circle}}, x_{\text{red}} x_{\text{square}} \rangle$ ("either a red circle or a red square").
- **Example:** $J = \langle x_{\text{red}} x_{\text{circle}}, x_{\text{red}} x_{\text{square}}, x_{\text{blue}} x_{\text{circle}}, x_{\text{blue}} x_{\text{square}} \rangle \subseteq R$.
  - We can "decompose" J as $J = \langle x_{\text{red}}, x_{\text{blue}} \rangle \cap \langle x_{\text{circle}}, x_{\text{square}} \rangle$.
  - Intuitively, $\langle x_{\text{red}}, x_{\text{blue}} \rangle$ and $\langle x_{\text{circle}}, x_{\text{square}} \rangle$ are concepts that are **simpler** than concept J.

Why define concepts as ideals?

1) To capture the expressiveness of human reasoning.

- We can construct **infinitely** many concepts from only finitely many primitive concepts. (See Theorem 3.1 of our paper.)
- The "richer" structure of ideals allows more operations, beyond set-theoretic operations. (See Section 4 of our paper.)
- In contrast to logic-based reasoning, we do **not** assign truth values (or numerical values) to variables. (See Appendix C.1 of our paper.)

2) For compatibility with *concept theory* in cognitive science.

- Compositional structure of concepts. (See Appendix A.4.2 of our paper.)
  - Important aspect of human experience in learning new concepts.
  - Concepts can be decomposed into intersections of ideals.
- Essence of concepts. (See Appendix A.4.2 of our paper.)
  - Concepts in terms of features: defining features vs irrelevant features.
  - Monomial ideals have computable unique "minimal generating sets".
- Concepts with partial definitions. (See Appendix A.4.2 of our paper.)
  - Humans can still reason with concepts that have partial definitions.
  - RPM task: Partial definition of $\langle x_{\text{square}} \rangle$ ("entity with four sides").

Attribute concepts

Idea: In human cognition, semantically similar concepts are naturally grouped to form a more general concept.

- **Example:** Concepts such as "red", "green", "blue", "yellow" can be grouped to form a new concept representing "color".
  - If $R = \mathbb{R}[x_{\text{red}}, x_{\text{green}}, x_{\text{blue}}, x_{\text{yellow}}]$, then the ideal $J \subseteq R$ given by $J = \langle x_{\text{red}}, x_{\text{green}}, x_{\text{blue}}, x_{\text{yellow}} \rangle$ is a concept that could mean "color".

Intuition: An attribute concept is a concept constructed by combining primitive concepts that represent "attribute values".

- **Example:** $\langle x_{\text{red}} \rangle$, $\langle x_{\text{green}} \rangle$, $\langle x_{\text{blue}} \rangle$, $\langle x_{\text{yellow}} \rangle$ are primitive concepts representing "attribute values" of the attribute concept J representing "color".
- To solve a reasoning task, we first need to identify *task-specific* attribute concepts.
  - RPM task: Attribute concepts "num", "pos", "type", "color", "size".

Inductive Bias: A concept representing a pattern shall be **deemed meaningful** if it is contained in some attribute concept.

What are primary decompositions?

Idea: Every ideal J has a **decomposition** $J = K_1 \cap \cdots \cap K_s$ as an intersection of finitely many primary ideals.

- This intersection is called a primary decomposition of J.
  - The constituent primary ideals K_i are called primary components.
  - If J is a monomial ideal, then there is a **unique** "minimal" primary decomposition $\mathrm{pd}(J)$ that we can compute.

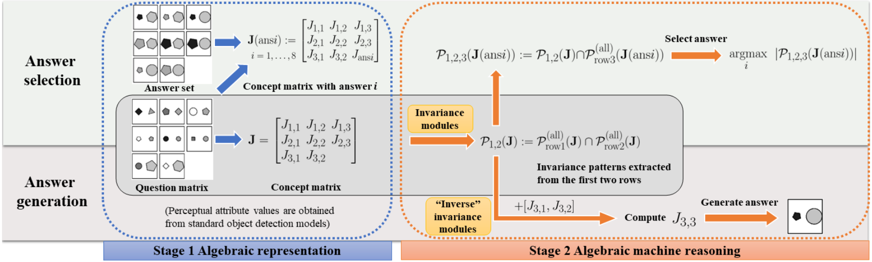
Reasoning via primary decompositions:

- A common pattern of concept $J_1, \ldots, J_k$ is a concept K containing $J_1, \ldots, J_k$.
  - If there are several common patterns $K_1, \ldots, K_r$, then we have:

$$\begin{aligned} J_1 &= K_1 \cap \cdots \cap K_r \cap K_1'; \\ J_2 &= K_1 \cap \cdots \cap K_r \cap K_2'; \\ &\vdots \\ J_k &= K_1 \cap \cdots \cap K_r \cap K_k'. \end{aligned}$$

- **Algebraic problem:** Compute common components $K_1, \ldots, K_r$.
  - Compute $\mathrm{pd}(J_1), \ldots, \mathrm{pd}(J_k)$, then extract the common primary components $K_1, \ldots, K_r$.
  - The new concept $K_1 \cap \cdots \cap K_r$ can be interpreted as a common pattern of $J_1, \ldots, J_k$.

Overview of algMR framework for solving RPMs



Perception Stage: Algebraic representation

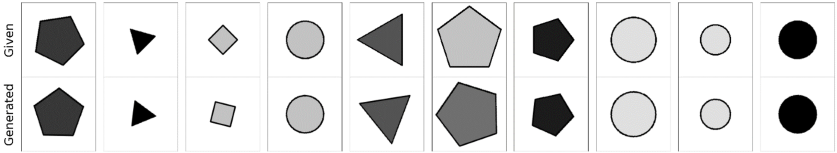
- Every (i, j)-th panel is represented as a concept $J_{i,j}$.
  - Raw perceptual attribute values are obtained from object detection models.
  - Each entity is encoded as a generator of $J_{i,j}$ (e.g. $x_{\text{small}} x_{\text{black}} x_{\text{square}}$).

Reasoning Stage: Algebraic machine reasoning

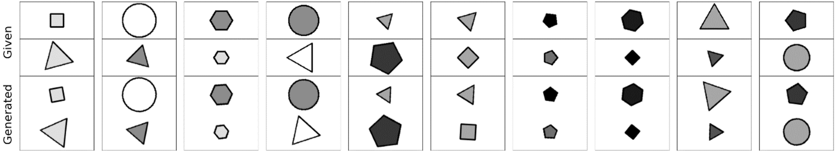
- Pattern extraction via four **invariance modules**.
  (1) intra-invariance module, (2) inter-invariance module,
  (3) compositional invariance module, (4) binary-operator invariance module
- **Intuition:** They check for 4 general types of invariances across a sequence of concepts.
  - Based on computing $\text{pd}(J_{i,j})$ and solving ideal-based problems.

Generative reasoning for RPMs

Our framework can generate answers with just the question matrix!



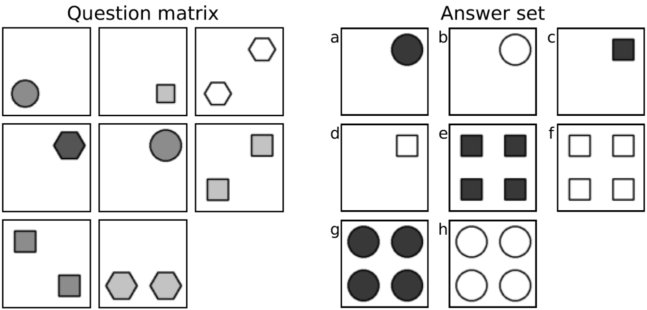
Top row: Given answers of randomly selected RPM instances.
Bottom row: Generated answers of respective RPM instances.



(See Section 3.4.2 of our paper for more details.)

- **Intuition:** Patterns extracted via the invariance modules can be used to "inversely" compute what the "correct answer" is.
  - Our generated answers achieved avg. similarity score of 67.7%.

RPM task: Discovering new patterns



Example: An RPM with an **unexpected new pattern discovered**.

- Our algMR framework selects both *b* and *h* as equally valid answers.
  - Given answer is option *h*.
    - Row-wise: Number of entities in first 2 panels sum up to number of entities in 3rd panel.
  - Option *b* is also valid!
    - **New discovered pattern:** Every panel has either 1 or 2 entities.

Beyond RPMs: Other reasoning tasks

Algebraic machine reasoning is well-suited for inductive reasoning.

- **Generalizable Idea:** "Discovering patterns" is realized concretely as "computing primary decompositions".
- Can be easily combined with any perception models.
  - Primitive concepts depend solely on what the machine can perceive.
    - e.g. object classes of its object detection models.
  - Perception need not be visual!
    - Can be tokens/words "perceived" by LLMs.
    - Can be abstract: e.g. robot's representation of environment.

More complicated reasoning tasks (e.g. automated theorem proving) require a mix of crystallized intelligence and fluid intelligence.

- We have LLMs for crystallized intelligence!
  - Even better if combined with knowledge bases!
- We have logic programming (LP) and its various extensions!
- Exciting possibilities: **LLMs + databases + LP + algMR**

Explainable Reasoning: Entire reasoning process is verifiable.

- Intermediate reasoning steps directly from algebraic computations.
  - NOTE: algMR is inherently **not search-based**.

Concepts: Beyond ideals of polynomial rings

Given an ideal I of the polynomial ring $S = \mathbb{R}[x_1, \ldots, x_n]$, we can construct a quotient ring $R = S/I$.

- The elements of the quotient ring R are **cosets** of I.
  - A coset of I is a subset of S of the form $p + I := \{p + q | q \in I\}$ for some polynomial $p \in S$.
  - This polynomial p is called a coset representative of $p + I$.
  - Convention: We write this coset simply as $\overline{p}$.
- **Example:** If $R = \mathbb{R}[x_1, x_2]/\langle x_1 - x_2 \rangle$, then $\overline{x_1}$ and $\overline{x_2}$ are the exact same element in R.
- A subset $J \subseteq R$ is called an ideal of R if there are cosets $g_1, \ldots, g_k$ in R such that

$$J = \{f_1 g_1 + \cdots + f_k g_k | f_1, \ldots, f_k \in R\},$$

  - We say that $\mathcal{G} = \{g_1, \ldots, g_k\}$ is a generating set for J.
  - If $g_1 = \overline{m_1}, \ldots, g_k = \overline{m_k}$ for some monomials $m_1, \ldots, m_k$ in S, then we say that J is a monomial ideal.

Extended Definition: A concept of R is a monomial ideal of R.

Final Remarks

Current state: Lots of work in intersection of automated reasoning and computational algebra (think: Gröbner bases and SMT).

With algMR, we have established a **new connection** between machine reasoning and commutative algebra (think: ideals as objects).

- Over a century's worth of deep results in commutative algebra.
  - Lots of algebraic operations on ideals we have not used.
    - e.g. ideal quotients, radicals, saturations, symbolic powers, etc.
  - Lots of non-search-based algebraic algorithms.
  - **Question:** Can we use them to extend reasoning capabilities?

So far: Concepts are monomial ideals of polynomial/quotient rings.

- monomial ideals $\rightsquigarrow$ arbitrary ideals.
- Coefficient field $\mathbb{R} \rightsquigarrow$ any Noetherian ring.

Know someone who might want to join the team?

- We are looking for research assistants, PhD students, and post-docs who are passionate about **algebraic methods in AI**.
  - Those with strong algebraic backgrounds are prioritized.
  - If you are interested, please directly contact Ernest Chong at `ernest_chong@sutd.edu.sg`

Thank you!


Jingyi Xu


Tushar Vaidya


Yufei Wu


Saket Chandra


Zhangsheng Lai


Ernest Chong