Analogical Proportions: What can you do with them in AI?

Mena Leemhuis
Free University of Bozen-Bolzano

Workshop "Analogies in Physics and Beyond"
November 26, 2024

AUTONOME PROVINZ BOZEN SÜDTIROL

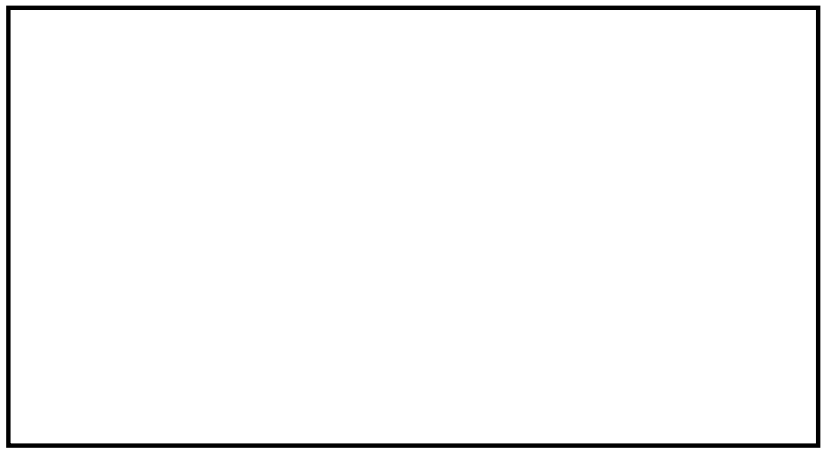
PROVINCIA AUTONOMA DI BOLZANO ALTO ADIGE

unibz

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances

Example (TransE[1])

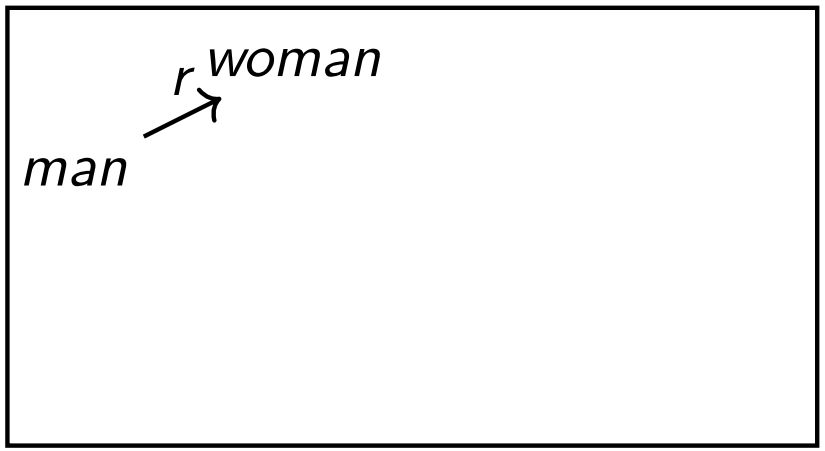


[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances

Example (TransE[1])

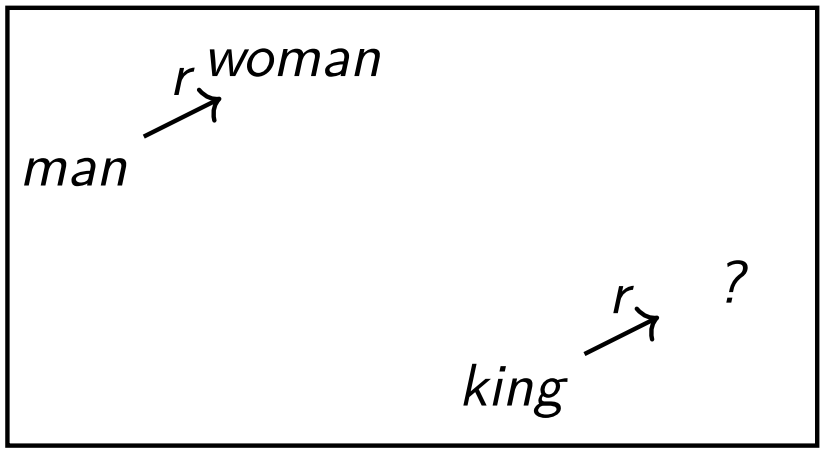


[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances
- question: *king* is male form of *?*

Example (TransE[1])

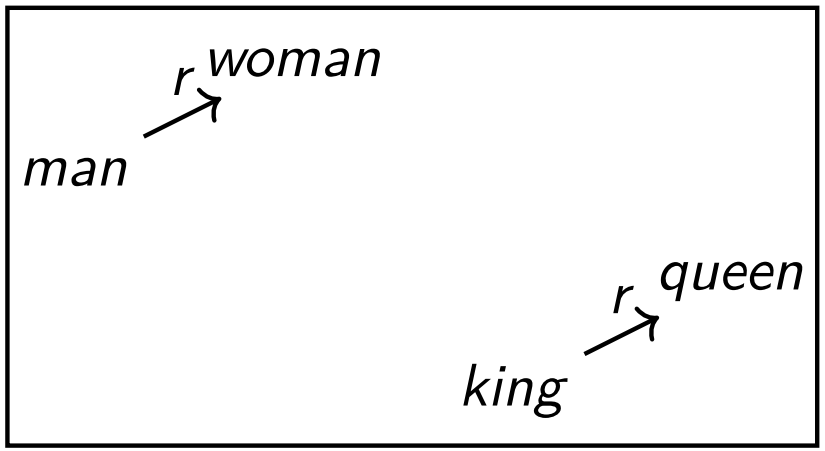


[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances
- question: *king* is male form of *?*

Example (TransE[1])

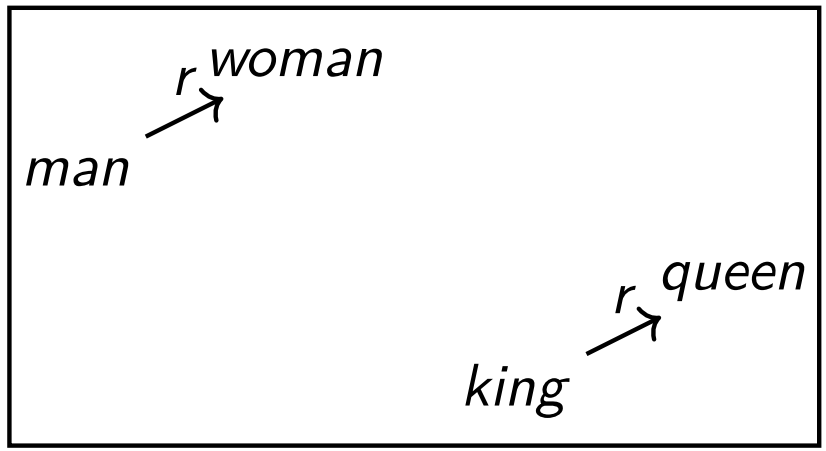


[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances
- question: *king* is male form of *?*
- man - male form of - woman leads to king male form of queen
- other writing: "man is to woman as king is to queen"

Example (TransE[1])



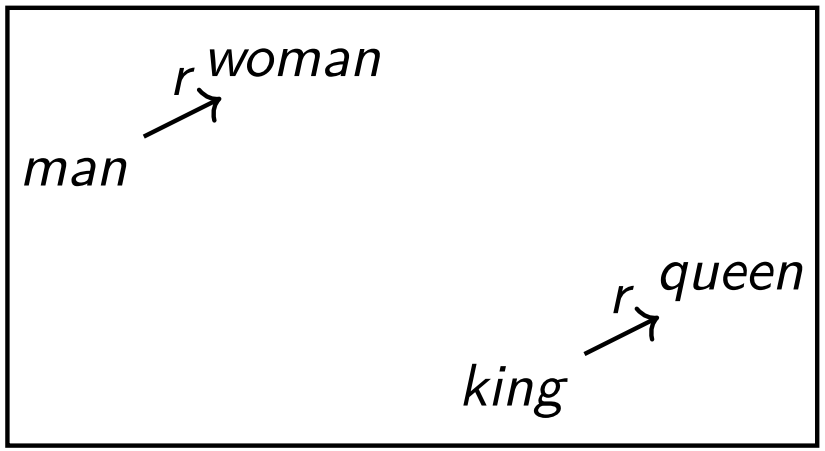
[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

Motivational example – knowledge graph embedding

- task: having instances and relations given, predicting new relations (links) between instances
- question: *king* is male form of *?*
- man - male form of - woman leads to king male form of queen
- other writing: "man is to woman as king is to queen"

→ analogical proportions (APs)

Example (TransE[1])



[1] Antoine Bordes et al. (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795

What is an analogical proportion?

- e.g., stated by Kant[2]

[2] Immanuel Kant (2004). *Immanuel Kant: Prolegomena to Any Future Metaphysics*. Ed. by Gary Hatfield. 2nd ed. Cambridge University Press.

What is an analogical proportion?

- e.g., stated by Kant[2]

  [Analogy] surely ***does not signify****, as the word is usually taken,* ***an imperfect similarity between two things****,*

[2] Immanuel Kant (2004). *Immanuel Kant: Prolegomena to Any Future Metaphysics*. Ed. by Gary Hatfield. 2nd ed. Cambridge University Press.

What is an analogical proportion?

- e.g., stated by Kant[2]

  [Analogy] surely ***does not signify****, as the word is usually taken,* ***an imperfect similarity between two things****,*

  but rather ***a perfect similarity between two relations in wholly dissimilar things***

[2] Immanuel Kant (2004). *Immanuel Kant: Prolegomena to Any Future Metaphysics*. Ed. by Gary Hatfield. 2nd ed. Cambridge University Press.

What is an analogical proportion?

- e.g., stated by Kant[2]

  [Analogy] surely ***does not signify****, as the word is usually taken,* ***an imperfect similarity between two things****,*

  but rather ***a perfect similarity between two relations in wholly dissimilar things***

- Therefore:

$$a \text{ is to } b \text{ as } c \text{ is to } d$$

[2] Immanuel Kant (2004). *Immanuel Kant: Prolegomena to Any Future Metaphysics*. Ed. by Gary Hatfield. 2nd ed. Cambridge University Press.

Motivational example – knowledge graph embedding

- *man* : *woman* :: *king* : *queen*
- *analogical ratio operator* : , *analogical proportion operator* ::

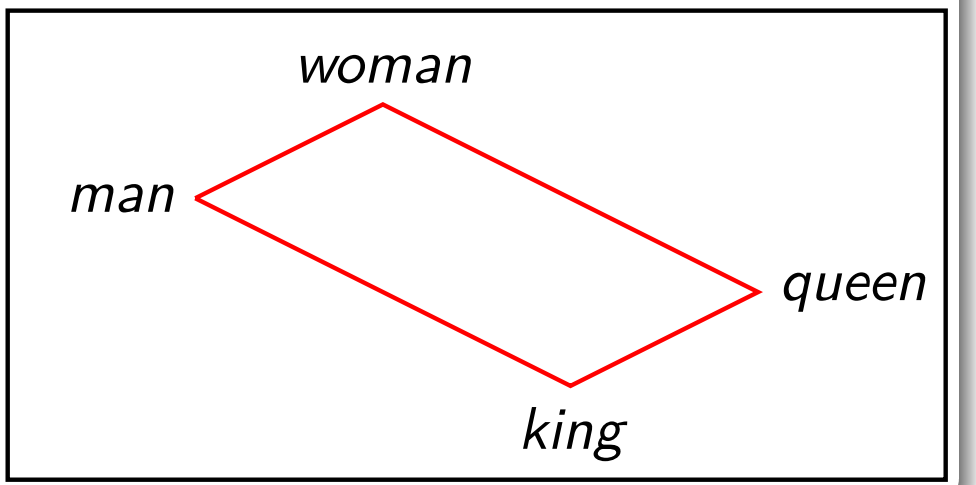
- also writable as

$$a - b = c - d$$

→ *arithmetic proportion*

- resembles classical view on APs (the parallelogram) [3]

Example (TransE)



[3] David E Rumelhart and Adele A Abrahamson (1973). "A model for analogical reasoning". In: *Cognitive Psychology* 5.1, pp. 1–28

A brief history of analogical proportions

- idea of APs goes back to Aristotle[4]

  [...] as old age is to life, so is evening to day. Evening may therefore be called "the old age of the day", and old age, "the evening of life" [...]

- several usage areas
  - language[5]:

    $$look : looked = walk : x \Rightarrow x = walked$$

  - arithmetic proportion: $a - b = c - d$
  - geometric proportion: $a/b = c/d$

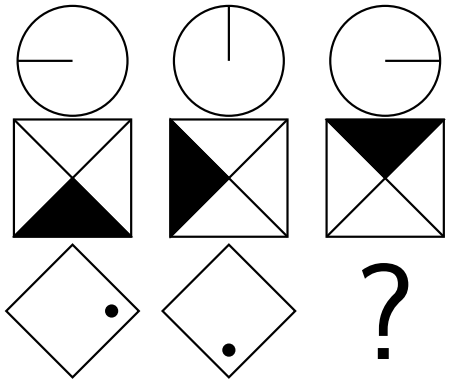
[4] Aristotle (1996). *Poetics*. Penguin Classics.

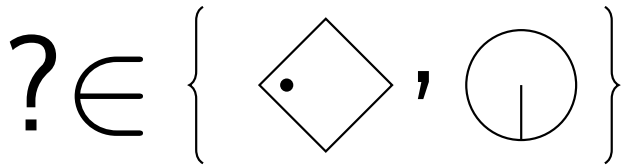
[5] Yves Lepage (2002). "Analogy and Formal Languages". In: *Electronic Notes in Theoretical Computer Science* 53, pp. 180–191.

Some examples for APs throughout AI

Ravens Progressive Matrices[6] (RPMs)

- non-verbal test for measuring general human intelligence
- creative or non-creative

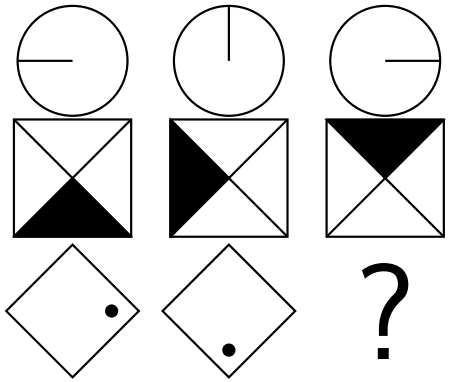


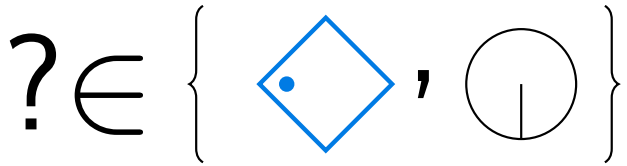


[6] John C. Raven (1941). "Standardzation of progressive matrices". In: *British Journal of Medical Psychology* 19.1, pp. 137–150.

Ravens Progressive Matrices[6] (RPMs)

- non-verbal test for measuring general human intelligence
- creative or non-creative





[6] John C. Raven (1941). "Standardzation of progressive matrices". In: *British Journal of Medical Psychology* 19.1, pp. 137–150.

Image analogies

seminal paper by Hertzmann et al.[7]

→ led to the area of *style transfer*[8] and *image-to-image translation*[9]

[7] Aaron Hertzmann et al. (2001). "Image analogies". In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*.

[8] Leon A. Gatys et al. (2016). "Image Style Transfer Using Convolutional Neural Networks". In: *2016 CVPR*, pp. 2414–2423.

[9] Phillip Isola et al. (2017). "Image-to-Image Translation with Conditional Adversarial Networks". In: *2017 CVPR*, pp. 5967–5976.

Miller Analogies Test

- *Carnivore* : *Herbivore* :: *Tiger* : (a. *lion*, b. *vegetation*)

Miller Analogies Test

- *Carnivore* : *Herbivore* :: *Tiger* : (a. *lion*, b. *vegetation*)

Miller Analogies Test

- *Carnivore* : *Herbivore* :: *Tiger* : (a. *lion*, b. *vegetation*)

All these approaches share the same question regarding APs: for given a, b, c determine d such that $a : b :: c : d$.

Miller Analogies Test

- *Carnivore* : *Herbivore* :: *Tiger* : (a. *lion*, b. *vegetation*)

All these approaches share the same question regarding APs: for given a, b, c determine d such that $a : b :: c : d$.

- analogously other positions of the missing element possible, e.g., *Poem* : (a. *line*, b. *rhyme*, c. *stanza*, d. *sonnet*) :: *Book* : *Chapter*

Miller Analogies Test

- *Carnivore* : *Herbivore* :: *Tiger* : (a. *lion*, b. *vegetation*)

All these approaches share the same question regarding APs: for given a, b, c determine d such that $a : b :: c : d$.

- analogously other positions of the missing element possible, e.g., *Poem* : (a. *line*, b. *rhyme*, c. *stanza*, d. *sonnet*) :: *Book* : *Chapter*

Analogical interpolation

- calculation of the "mean" with *continuous analogical proportions* ($a : x :: x : c$ with missing x)
  - e.g., for the arithmetic proportion:

$$a : x :: x : c \rightarrow a - x = x - c \rightarrow a + c = 2x \rightarrow \frac{a+c}{2} = x$$

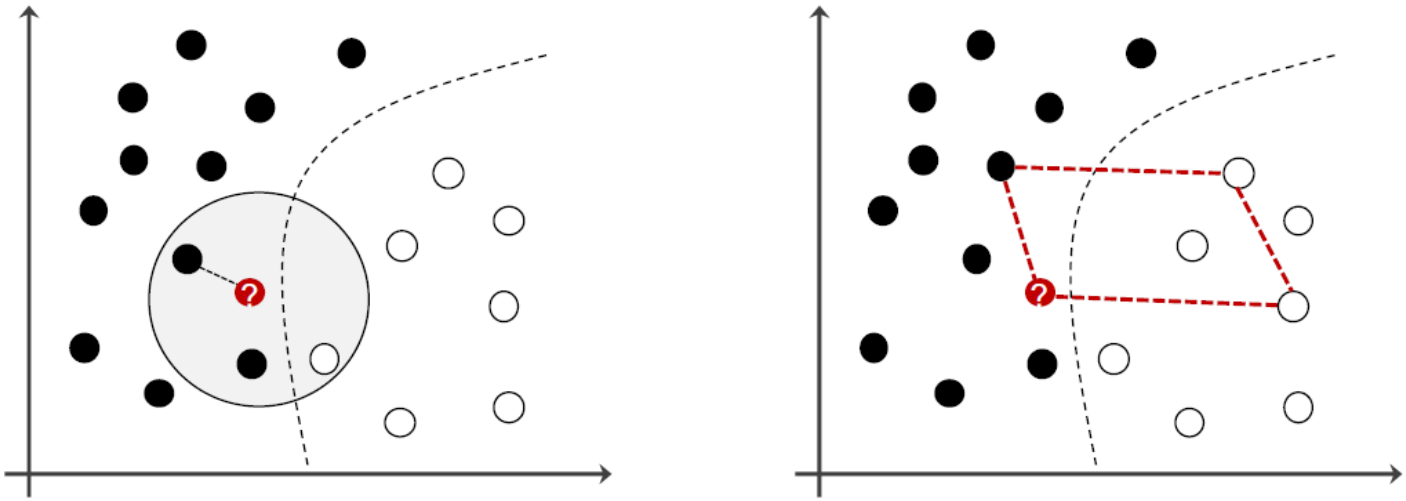
  - creation of an "intermediate" item in between two others, e.g.,

$$\textit{man} : \textit{centaur} :: \textit{centaur} : \textit{horse}$$

- usable for enlarging datasets (see, e.g., Bounhas and Prade[10])

[10] Myriam Bounhas and Henri Prade (2019). "An analogical interpolation method for enlarging a training dataset". In: *Scalable Uncertainty Management: 13th International Conference, SUM 2019, Compiègne, France, December 16–18, 2019, Proceedings 13*. Springer, pp. 136–152.

Analogy-based explanations



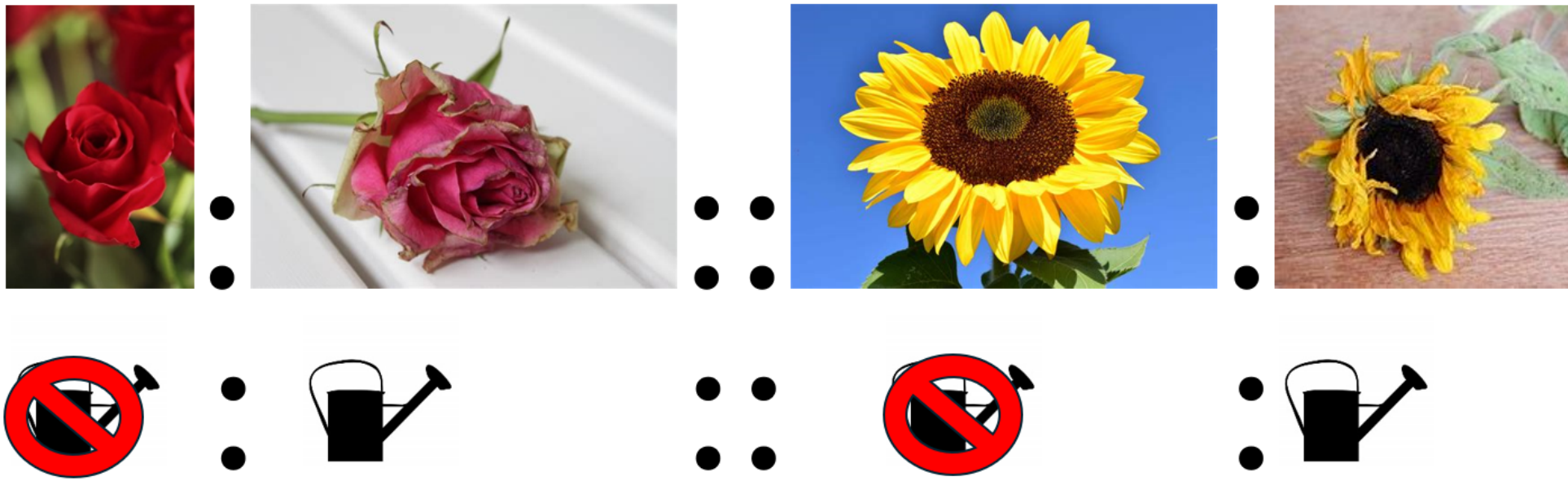
- explanation of classification results with the help of APs by Hüllermeier[11]
- explanation based on similar elements could be misleading
- explanation with AP allows for applying the relation between known elements to the relation between a known and an unknown one

[11] Eyke Hüllermeier (2020). "Towards Analogy-Based Explanations in Machine Learning". In: *Modeling Decisions for Artificial Intelligence*, pp. 205–217.

Case-based reasoning

- proposed, e.g., by Lieber et al.[12]:
- given four problems p_a, p_b, p_c, p_d and three solutions s_a, s_b, s_c:

  Knowing that $p_a : p_b :: p_c : p_d$ leads to the assumption of $s_a : s_b :: s_c : s_d$ and enables finding an s_d



[12] Jean Lieber et al. (2018). "Making the Best of Cases by Approximation, Interpolation and Extrapolation". In: *Case-Based Reasoning Research and Development*, pp. 580–596.

What is an Analogical Proportion?

What is an AP?

- Quaternary relation of the form

  "*a* is to *b* as *c* is to *d*" (R1)

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "a is to b as c is to d" (R1)

- Formal notation "$a : b :: c : d$" suggests other reading

  "Ratio of a and b is similar to ratio of c and d" (R2)

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "a is to b as c is to d" (R1)

- Formal notation "$a : b :: c : d$" suggests other reading

  "Ratio of a and b is similar to ratio of c and d" (R2)

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "a is to b as c is to d" (R1)

- Formal notation "$a : b :: c : d$" suggests other reading

  "Ratio of a and b is similar to ratio of c and d" (R2)

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "*a* is to *b* as *c* is to *d*" (R1)

- Formal notation "*a* : *b* :: *c* : *d*" suggests other reading

  "Ratio of *a* and *b* is similar to ratio of *c* and *d*" (R2)

Challenge

How do we get from (R1) to (R2)?

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "*a* is to *b* as *c* is to *d*" (R1)

- Formal notation "*a* : *b* :: *c* : *d*" suggests other reading

  "Ratio of *a* and *b* is similar to ratio of *c* and *d*" (R2)

Challenge

How do we get from (R1) to (R2)?

- Instance of *extension problem for compositionality of the first form*[13]

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

What is an AP?

- Quaternary relation of the form

  "*a* is to *b* as *c* is to *d*" (R1)

- Formal notation "*a* : *b* :: *c* : *d*" suggests other reading

  "Ratio of *a* and *b* is similar to ratio of *c* and *d*" (R2)

Challenge

How do we get from (R1) to (R2)?

- Instance of *extension problem for compositionality of the first form*[13]

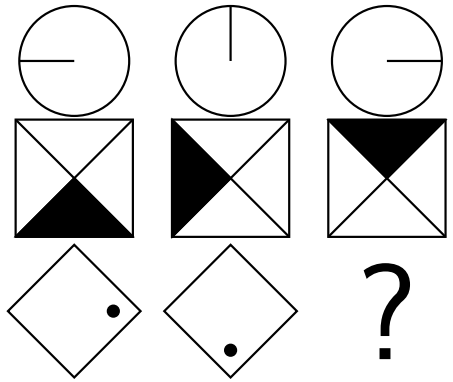
Challenge

How do we get interpretable/explainable APs?

[13] Wilfrid Hodges (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

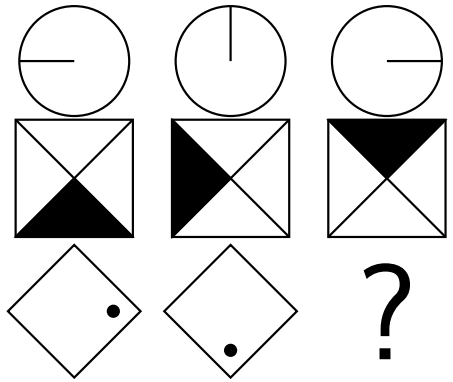
Example – How to solve Raven's Progressive Matrices?

1. How could the problem be syntactically interpreted as AP?
2. Is the problem actually an AP?
3. How to solve the AP?



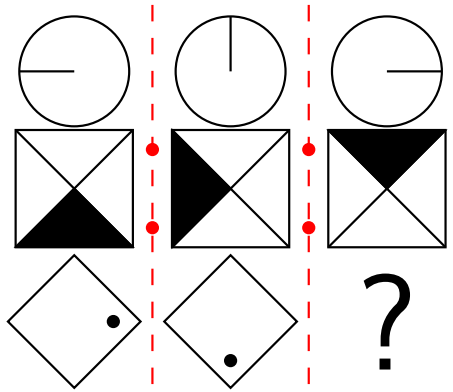
RPMs as analogical proportions (non-uniqueness – part I)

- RPMs can be interpreted as APs (however, this is not always the case)
- definition is not unique



RPMs as analogical proportions (non-uniqueness – part I)

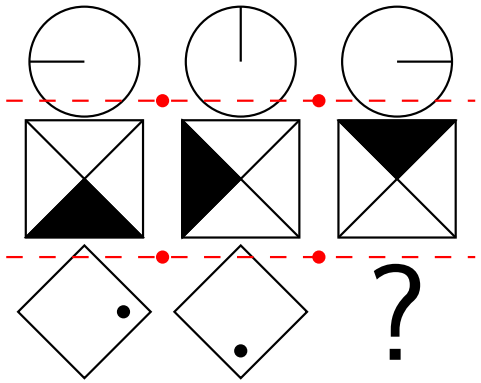
- RPMs can be interpreted as APs (however, this is not always the case)
- definition is not unique



- Analogical ratio between columns

RPMs as analogical proportions (non-uniqueness – part I)

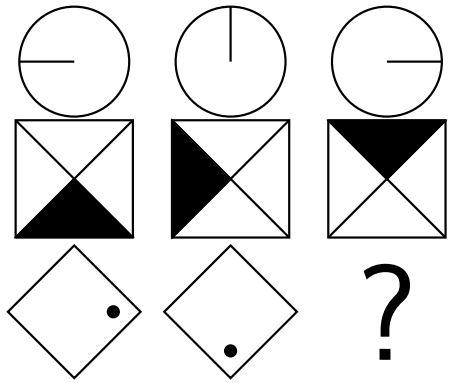
- RPMs can be interpreted as APs (however, this is not always the case)
- definition is not unique



- Analogical ratio between rows

RPMs as analogical proportions (non-uniqueness – part I)

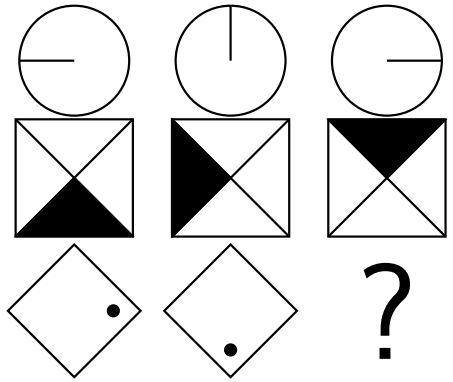
- RPMs can be interpreted as APs (however, this is not always the case)
- definition is not unique



- Analogical ratio between figures

RPMs as analogical proportions (non-uniqueness – part I)

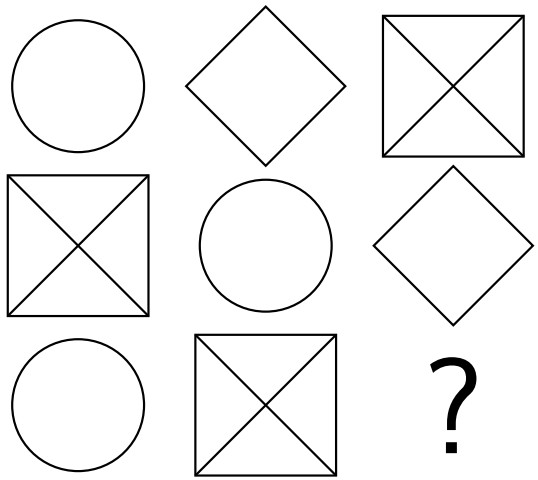
- RPMs can be interpreted as APs (however, this is not always the case)
- definition is not unique



- or even:
  $(p_1, p_2) : p_3 :: (q_1, q_2) : q_3$

Are RPMs actually analogy problems?

- example dataset: RAVEN[14]
- includes, e.g., "distribute three"



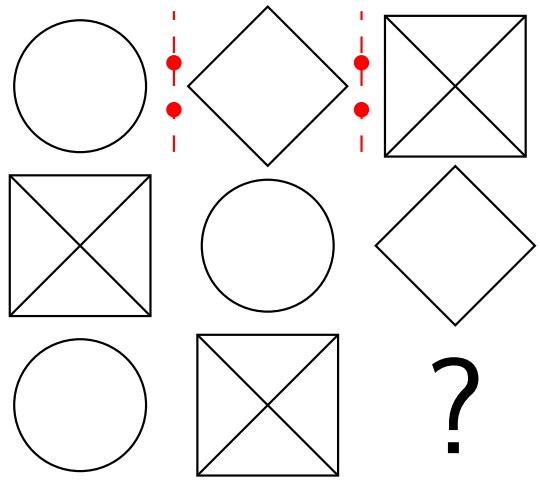
Problem

Not everything based on four components is an AP (or at least not in every interpretation).

[14] Chi Zhang et al. (2019). "RAVEN: A Dataset for Relational and Analogical Visual REasoNing". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5312–5322

Are RPMs actually analogy problems?

- example dataset: RAVEN[14]
- includes, e.g., "distribute three"



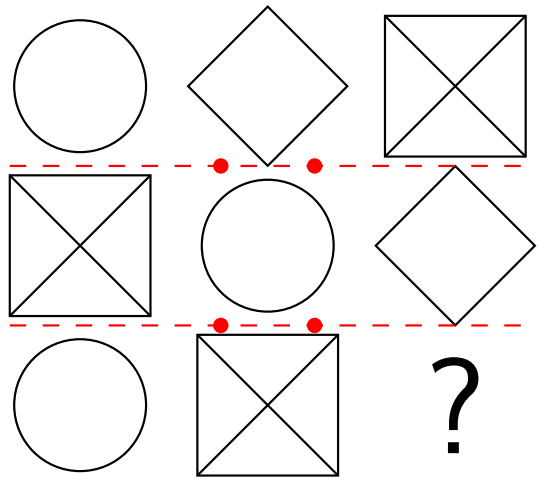
Problem

Not everything based on four components is an AP (or at least not in every interpretation).

[14] Chi Zhang et al. (2019). "RAVEN: A Dataset for Relational and Analogical Visual REasoNing". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5312–5322

Are RPMs actually analogy problems?

- example dataset: RAVEN[14]
- includes, e.g., "distribute three"



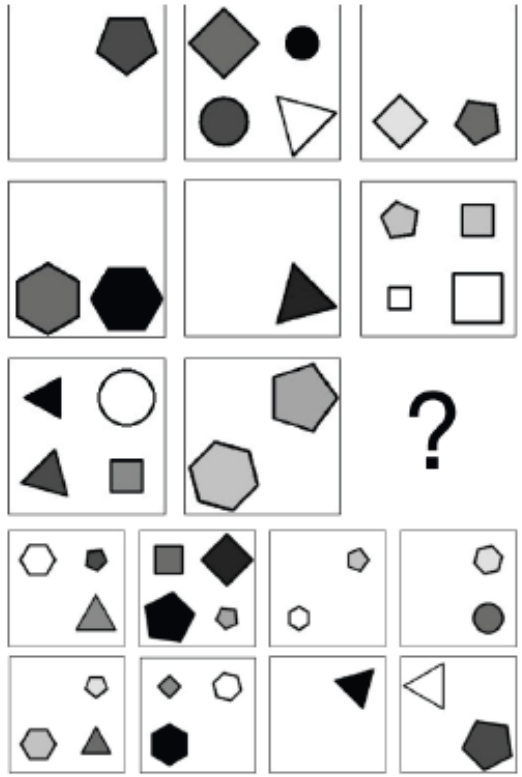
Problem

Not everything based on four components is an AP (or at least not in every interpretation).

[14] Chi Zhang et al. (2019). "RAVEN: A Dataset for Relational and Analogical Visual REasoNing". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5312–5322

How to find the best solution (non-uniqueness – part II)

- Which AP is the best one?
- Is there even a single best one?



(from Zhang et al., 2019)

Challenge

How to make sure the "best" AP is found?

Naive approaches for solving RPMs

e.g., by Zhuo and Kankanhalli[15]

- neural network with incomplete RPM and choices for missing images as input
- supervised learning based on known analogical proportions as training data
- leads to acceptable result quality

Problem

- *no decomposition*
- *no explainability*
- *not even any underlying structure*

$\rightarrow$ *no usage of analogical proportions!*

[15] Tao Zhuo and Mohan S. Kankanhalli (2020). "Solving Raven's Progressive Matrices with Neural Networks". In: *ArXiv* abs/2002.01646.

Challenge

A structured definition of APs is needed

- for numbers, the arithmetic or geometric proportion can be used (thus, e.g., $a/b = c/d$)

$\rightarrow$ How to adapt this structure to the general case of APs between arbitrary elements?

The basic formalism

Definition (Basic axioms (e.g., by Prade and Richard[16]))

For a set of items X, $\forall a, b, c, d \in X$:

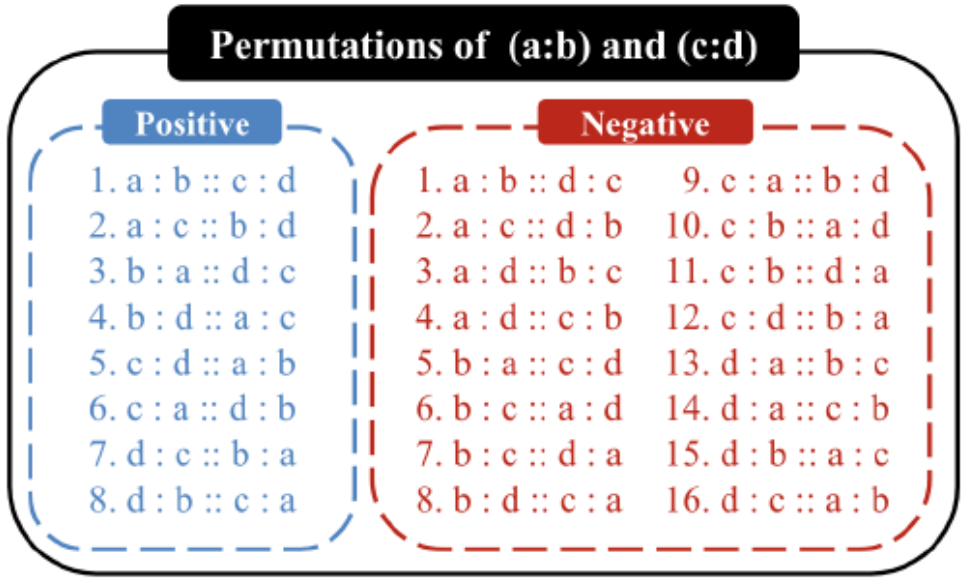
- $a : b :: a : b$ (reflexivity)
- $a : b :: c : d \rightarrow c : d :: a : b$ (symmetry)
- $a : b :: c : d \rightarrow a : c :: b : d$ (central permutation)

[16] Henri Prade and Gilles Richard (2014). "From Analogical Proportion to Logical Proportions: A Survey". In: *Studies in Computational Intelligence*, pp. 217–244

An approach incorporating basic axioms

presented by Ushio et al.[17]

- not based on RPMs, however, idea easily adaptable
- incorporation of basic axioms in the loss function
- leads to more explicit inspection of APs
- however:
  - no explainability
  - only a slight restriction of possible proportions
  - no explicit meaning of : and ::



[17] Asahi Ushio et al. (2021). "BERT is to NLP what AlexNet is to CV: Can Pre-Trained Language Models Identify Analogies?" In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, pp. 3609–3624.

- Thus: even more structure needed
- First step: simplifying the problem

→ considering features

man

- male
- non-royal

woman

- female
- non-royal

king

- male
- royal

queen

- female
- royal

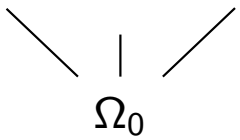
Back to the roots: the Boolean domain $\{0, 1\}$

- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]

[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

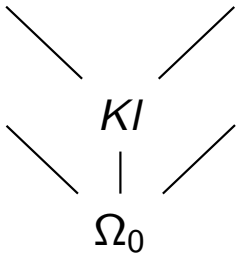
- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]
  - $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)



[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

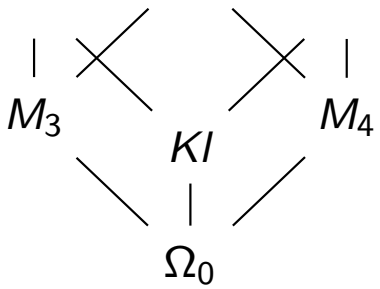
- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]
  - $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
  - $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)



[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]

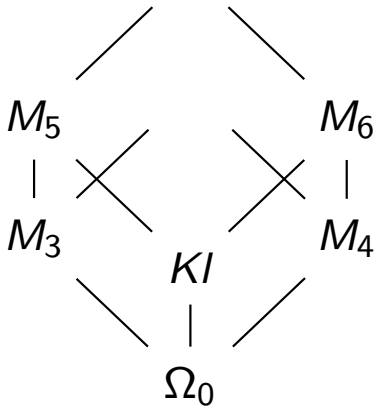


- $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
- $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)
- $M_3 = \Omega_0 \cup \{1110, 1101, 1011, 0111\}$
- $M_4 = \Omega_0 \cup \{0001, 0010, 0100, 1000\}$

[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]

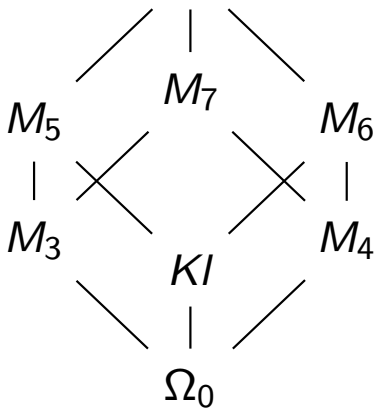


- $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
- $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)
- $M_3 = \Omega_0 \cup \{1110, 1101, 1011, 0111\}$
- $M_4 = \Omega_0 \cup \{0001, 0010, 0100, 1000\}$
- $M_5 = M_3 \cup \{0110, 1001\} = M_3 \cup KI$
- $M_6 = M_4 \cup \{0110, 1001\} = M_4 \cup KI$

[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]

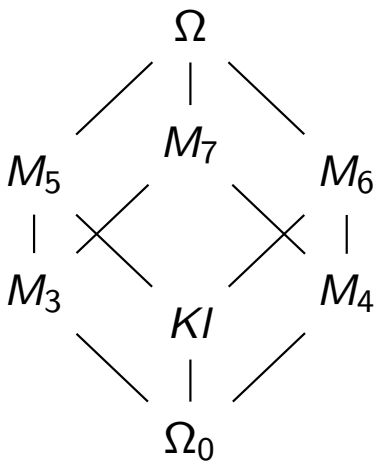


- $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
- $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)
- $M_3 = \Omega_0 \cup \{1110, 1101, 1011, 0111\}$
- $M_4 = \Omega_0 \cup \{0001, 0010, 0100, 1000\}$
- $M_5 = M_3 \cup \{0110, 1001\} = M_3 \cup KI$
- $M_6 = M_4 \cup \{0110, 1001\} = M_4 \cup KI$
- $M_7 = M_3 \cup \{0001, 0010, 0100, 1000\} = M_3 \cup M_4$

[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Back to the roots: the Boolean domain $\{0, 1\}$

- $a : b :: c : d$ becomes a quaternary Boolean function
- Each such function can be identified with the set of 4-bit vectors with evaluation 1
- Lattice of 8 models of analogical proportions by Prade and Richard[18]

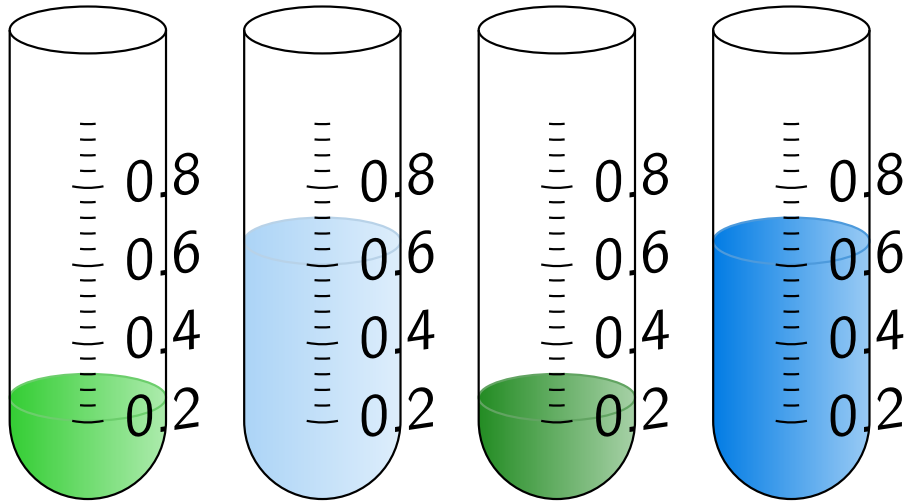


- $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
- $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)
- $M_3 = \Omega_0 \cup \{1110, 1101, 1011, 0111\}$
- $M_4 = \Omega_0 \cup \{0001, 0010, 0100, 1000\}$
- $M_5 = M_3 \cup \{0110, 1001\} = M_3 \cup KI$
- $M_6 = M_4 \cup \{0110, 1001\} = M_4 \cup KI$
- $M_7 = M_3 \cup \{0001, 0010, 0100, 1000\} = M_3 \cup M_4$
- $\Omega = \{0, 1\}^4 = \{0000, 0001, 0010, \ldots, 1111\}$

[18] Henri Prade and Gilles Richard (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.

Generalization part I – Fuzziness[19]

- $a, b, c, d \in [0, 1]$, therefore fuzzy

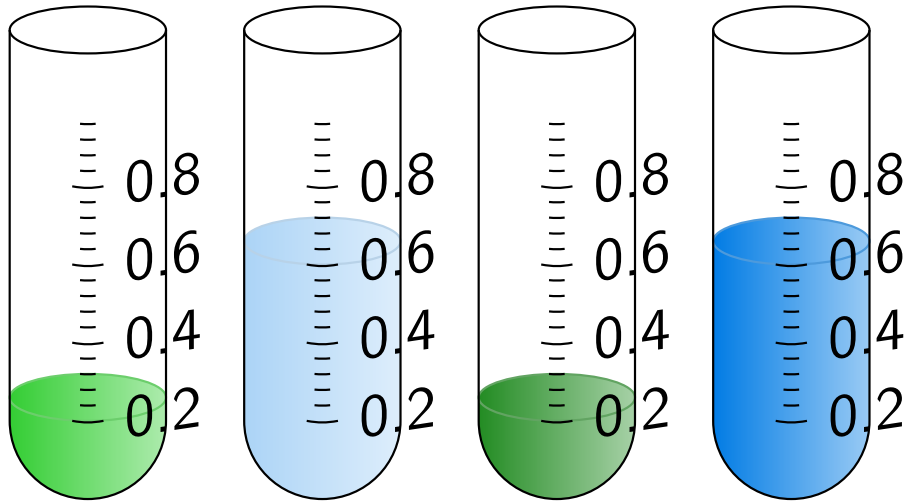


[19] Didier Dubois et al. (2016). "Multiple-valued extensions of analogical proportions". In: *Fuzzy Sets and Systems* 292, pp. 193–202.

Generalization part I – Fuzziness[19]

- $a, b, c, d \in [0, 1]$, therefore fuzzy
- either conservative, e.g.,

$$0.2 : 0.6 :: 0.2 : 0.6$$



[19] Didier Dubois et al. (2016). "Multiple-valued extensions of analogical proportions". In: *Fuzzy Sets and Systems* 292, pp. 193–202.

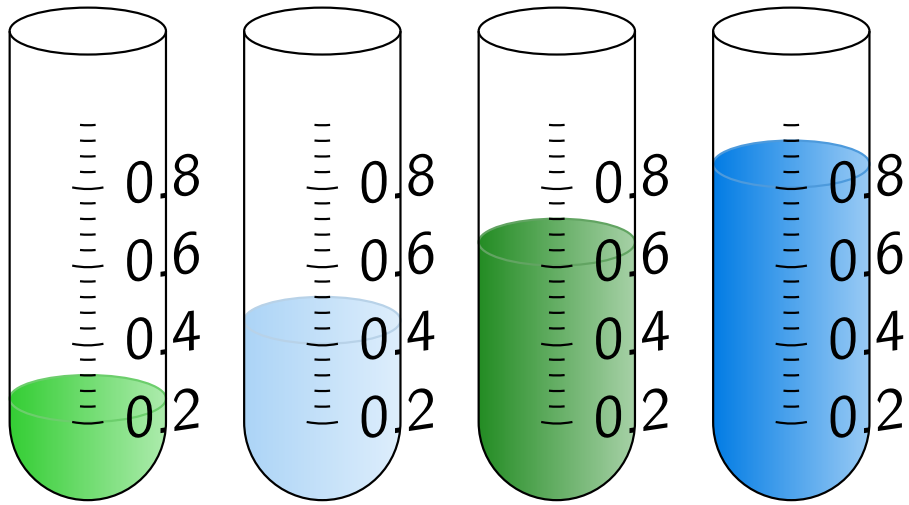
Generalization part I – Fuzziness[19]

- $a, b, c, d \in [0, 1]$, therefore fuzzy
- either conservative, e.g.,

$$0.2 : 0.6 :: 0.2 : 0.6$$

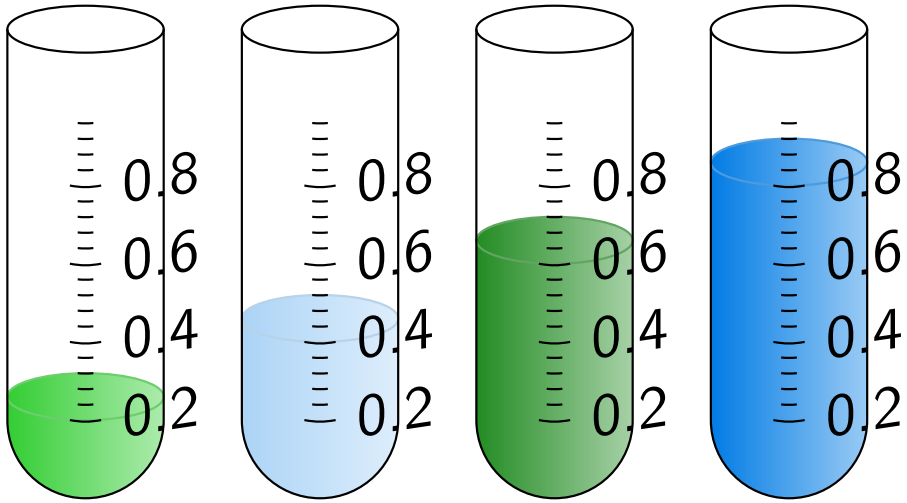
- or liberal graded, e.g.,

$$0.2 : 0.4 :: 0.6 : 0.8$$



[19] Didier Dubois et al. (2016). "Multiple-valued extensions of analogical proportions". In: *Fuzzy Sets and Systems* 292, pp. 193–202.

Generalization part II – Boolean vectors



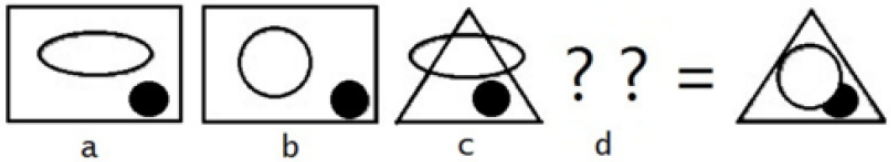
- component-wise extension from X to X^n
- $a : b :: c : d$ iff $\forall i \in \{1, \ldots, n\}, a_i : b_i :: c_i : d_i$
- also possible in the fuzzy case

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

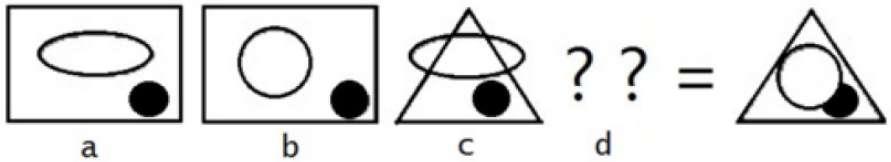
	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*?*	*?*	*?*	*?*	*?*

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

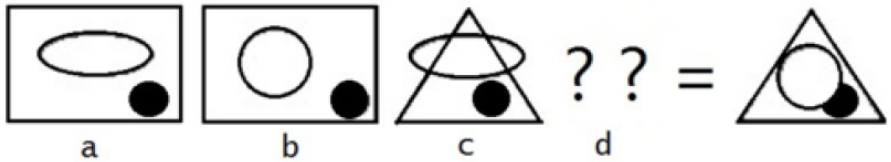
	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*0*	*?*	*?*	*?*	*?*

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

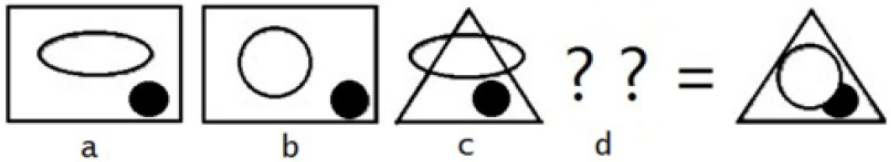
	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*0*	*1*	*?*	*?*	*?*

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

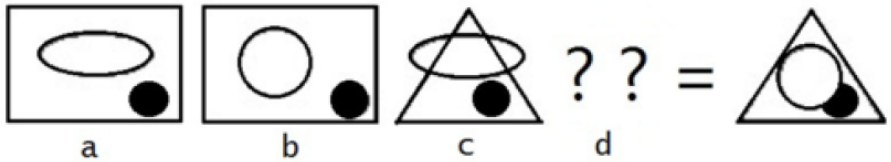
	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*0*	*1*	*1*	*?*	*?*

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

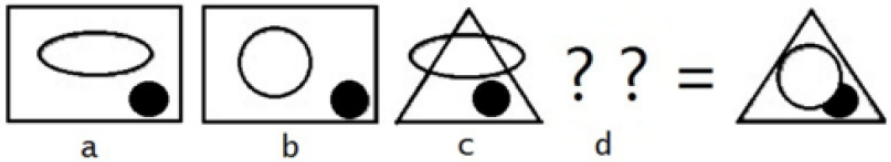
	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*0*	*1*	*1*	*1*	*?*

RPMs – straightforward solution

- based on feature vectors
- advantage:
  - explainability
- problems:
  - correct annotation of images needed
  - choice of features already incorporates implicit knowledge about analogy
  - normally, analogies are not optimal but noisy
    $\rightarrow$ *feature selection* necessary

$\rightarrow$ therefore, generalization necessary, abstraction from pure feature vectors

Example (from Correa et al., 2012)



- *hS: hasSquare*
- *hBD: hasBlackDot*
- *hT: hasTriangle*
- *hC: hasCircle*
- *hE: hasEllipse*

	hS	*hBD*	*hT*	*hC*	*hE*
a	*1*	*1*	*0*	*0*	*1*
b	*1*	*1*	*0*	*1*	*0*
c	*0*	*1*	*1*	*0*	*1*
d	*0*	*1*	*1*	*1*	*0*

Towards a (more general) generalization[20]

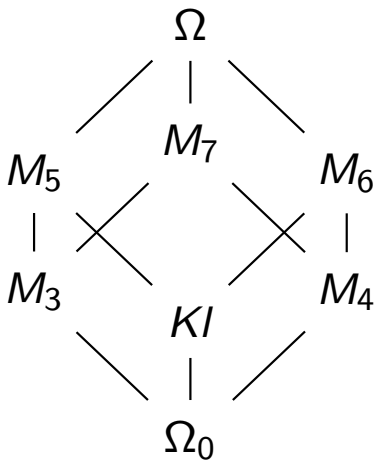
- one way of generalization: examining the underlying rules of each of the models

Formally:

- Want to find refined axiomatizations
- For each analogical proportion m find axiom set F_m such that
  1. $Mod(B_{ax} \cup F_m) = \{m\}$
  2. F_m is simple
  3. F_m is general

[20] Mena Leemhuis, Diedrich Wolter, et al. (2024). "Decomposing Analogy: A Logic Characterization". In: *Foundations of Information and Knowledge Systems*, pp. 256–274.

Refined axiomatization

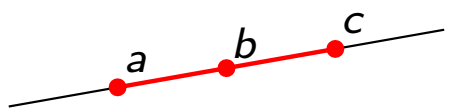


- $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$ (minimal model)
- $KI = \Omega_0 \cup \{0110, 1001\}$ (Klein's model)
- $M_3 = \Omega_0 \cup \{1110, 1101, 1011, 0111\}$
- $M_4 = \Omega_0 \cup \{0001, 0010, 0100, 1000\}$
- $M_5 = M_3 \cup \{0110, 1001\} = M_3 \cup KI$
- $M_6 = M_4 \cup \{0110, 1001\} = M_4 \cup KI$
- $M_7 = M_3 \cup \{0001, 0010, 0100, 1000\} = M_3 \cup M_4$
- $\Omega = \{0, 1\}^4 = \{0000, 0001, 0010, \ldots, 1111\}$

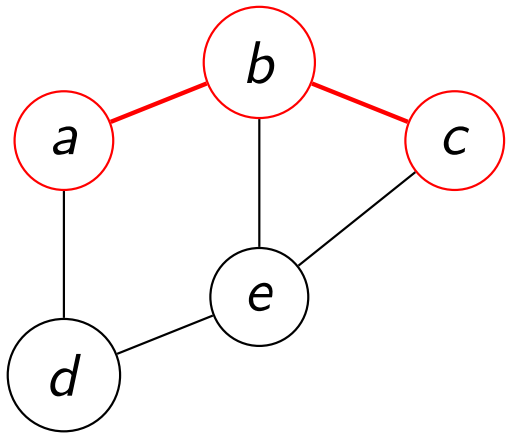
- $F_{\Omega_0} = \{unicity\} \cup \{antisymmetry\}$
- unicity: $\forall a \forall b \forall c \, (a : a :: b : c \rightarrow b = c)$
- antisymmetry: $\{\forall a \forall b \, (a : b :: b : a \rightarrow a = b)\}$

Betweenness-based view[21]

- creation of a generalized geometrical view of APs
- with the help of betweenness
- more general than arithmetic or geometric proportions



$Btw(a, b, c)$



[21] Mena Leemhuis and Özgür L. Özçep (2023). "Analogical Proportions and Betweenness". In: *Proceedings of the 9th Workshop on Formal and Cognitive Reasoning, co-located with the 46th German Conference on Artificial Intelligence (KI 2023)*, pp. 8–19.

Towards a (more general) generalization - variant II

- proposed by Herzig et al.[22](and within a different framework by Antić[23])
- considering APs between formulas with the help of transformation functions
- basic idea: considering three types of feature changes
  - true to false
  - false to true
  - irrelevant

man

- male
- ~~non-royal~~

woman

- female
- ~~non-royal~~

king

- male
- royal

queen

- female
- royal

[22] Andreas Herzig et al. (2024). "A Novel View of Analogical Proportion Between Formulas". In: *ECAI 2024*

[23] Christian Antić (2022). "Analogical proportions". In: *Annals of Mathematics and Artificial Intelligence* 90.6, pp. 595–644

Central permutation – intuition

- man:woman::king:queen (based on the relation *male form of*)
- man:king::woman:queen (based on the relation *royal form of*)

→ after central permutation, a completely new relation is needed

- But: problem with domain change:

$$F1\ \textit{driver} : \textit{car} :: \textit{jockey} : \textit{horse}$$

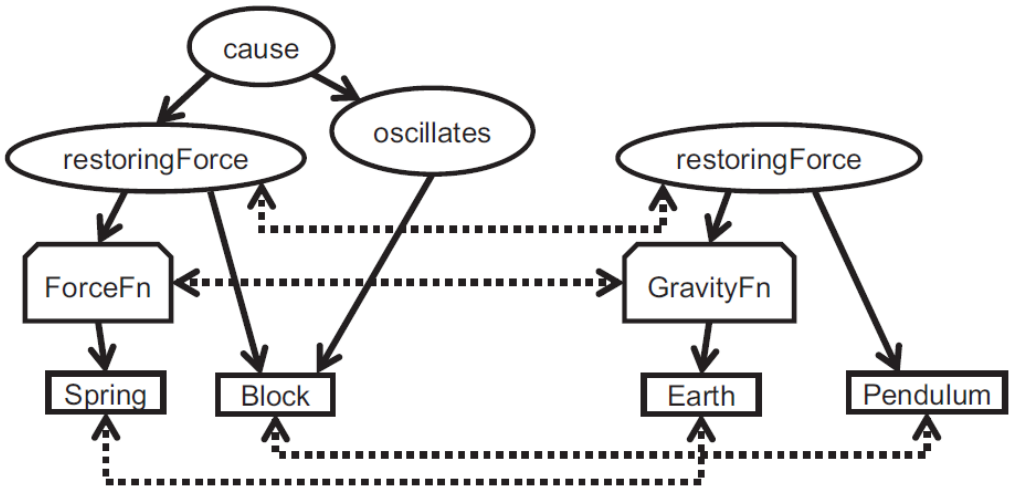
vs.

$$F1\ \textit{driver} : \textit{jockey} :: \textit{car} : \textit{horse}$$

- again, dependent on the framework considered

A different viewpoint – the Structure Mapping Engine

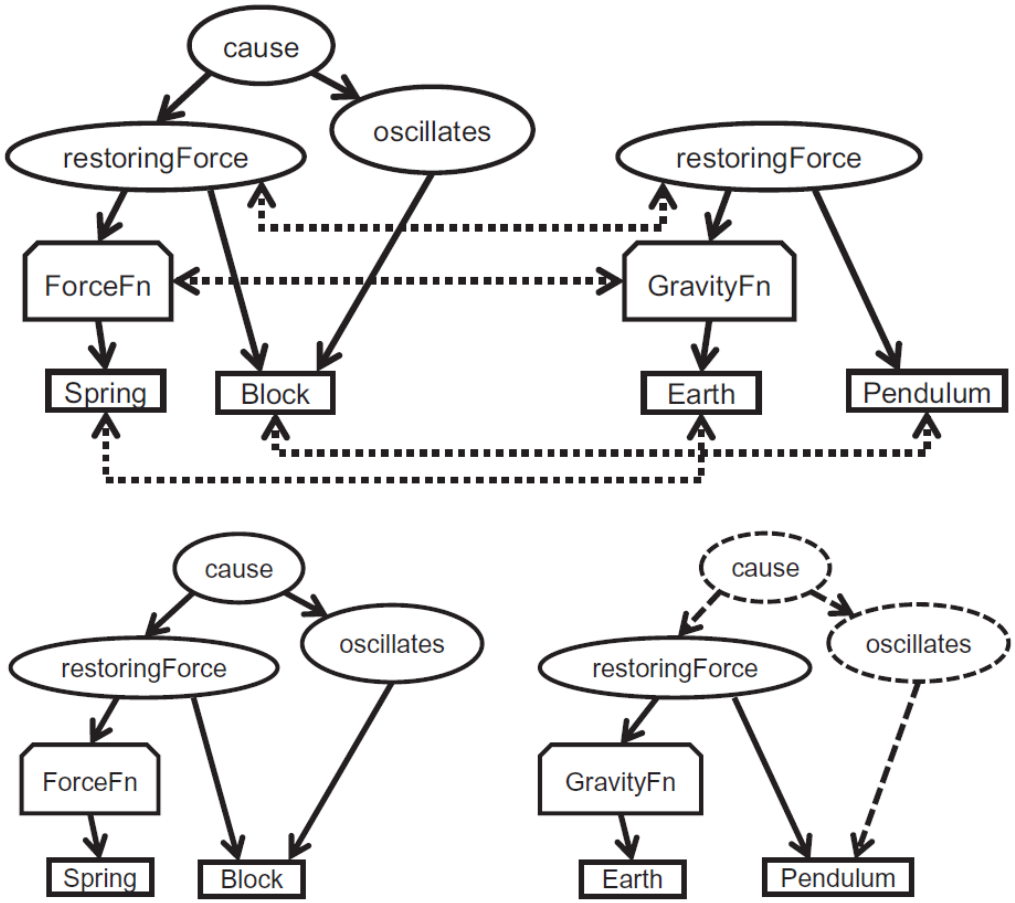
A different viewpoint – the Structure Mapping Engine[24]



(figures from Forbus, 2019)

[24] Brian Falkenhainer et al. (1989). "The structure-mapping engine: Algorithm and examples". In: *Artificial Intelligence* 41.1, pp. 1–63.

A different viewpoint – the Structure Mapping Engine[24]

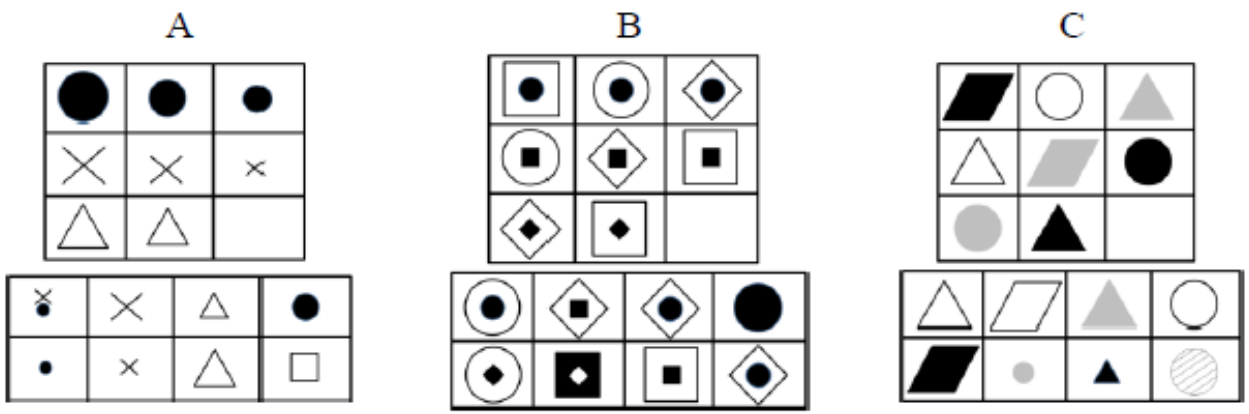


(figures from Forbus, 2019)

[24] Brian Falkenhainer et al. (1989). "The structure-mapping engine: Algorithm and examples". In: *Artificial Intelligence* 41.1, pp. 1–63.

Solving RPMs with the help of Structure Mapping

- presented by Lovett et al.[25]
- creation of pattern of variance based on a strategy for the top two rows
- then comparing a possible third row with the top two rows

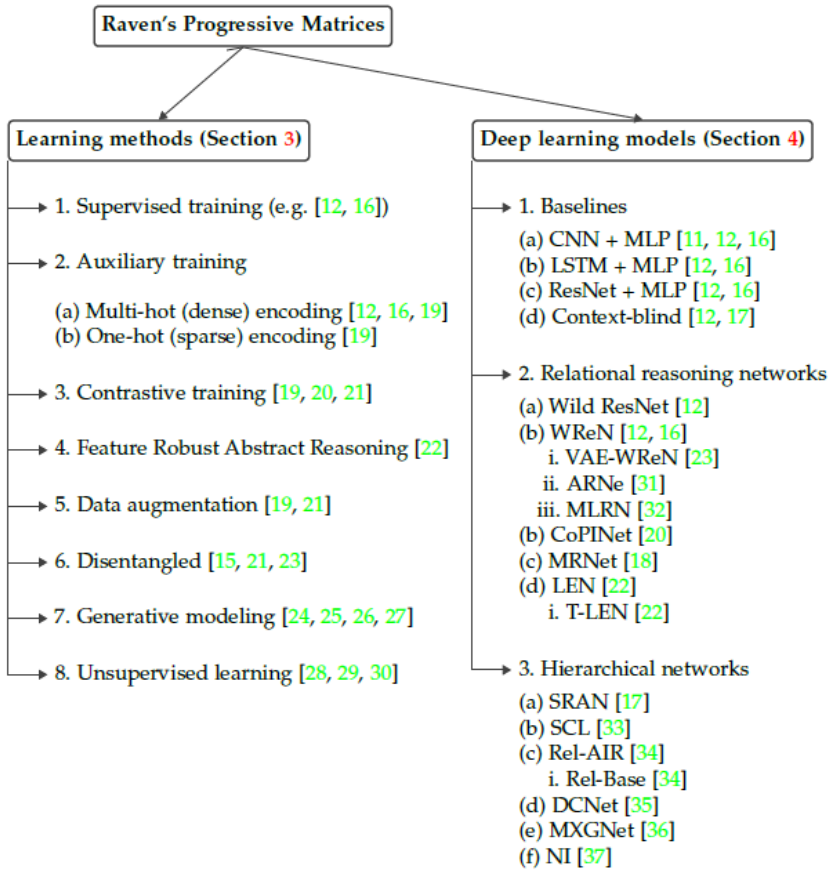


Carpenter Rules	Quantitative Pairwise Progression	Constant in a Row + Distribution of Three	Distribution of Three (applies twice)
Our Classification	Differences	Literal	Advanced Literal
Answer	3	5	2

[25] Andrew M. Lovett et al. (2010). "A Structure-Mapping Model of Raven's Progressive Matrices". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 32.

Solving RPMs – conclusion

- RPM's are heavily studied, not only in view of APs
- overview about state-of-the-art approaches, e.g., by Małkiński and Mańdziuk[26]



[26] Mikolaj Małkiński and Jacek Mańdziuk (2022). "Deep Learning Methods for Abstract Visual Reasoning: A Survey on Raven's Progressive Matrices". In: *ArXiv* abs/2201.12382

Does ChatGPT solve the issues with APs?

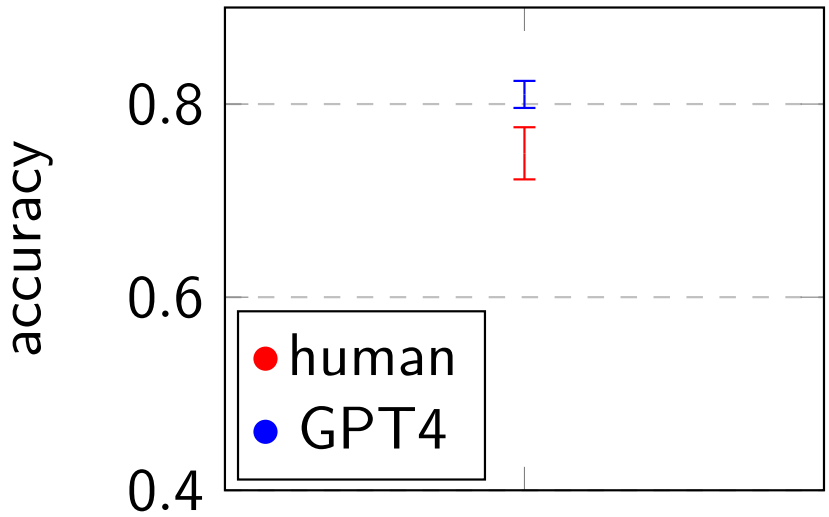
- Study by Webb et al.[27] about analogical reasoning capabilities of ChatGPT

[27] Taylor Webb et al. (2023). "Emergent analogical reasoning in large language models". In: *Nature Human Behaviour* 7.9, pp. 1526–1541

Does ChatGPT solve the issues with APs?

- Study by Webb et al.[27] about analogical reasoning capabilities of ChatGPT

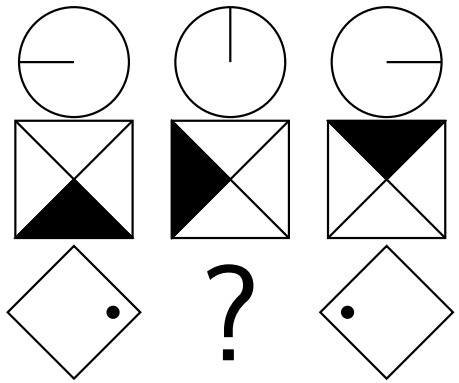
- results for a number-matrix problem (results from Mitchell, 2024)



[27] Taylor Webb et al. (2023). "Emergent analogical reasoning in large language models". In: *Nature Human Behaviour* 7.9, pp. 1526–1541

No, it doesn't

- different study on generalized analogical reasoning problems by Lewis and Mitchell[28]

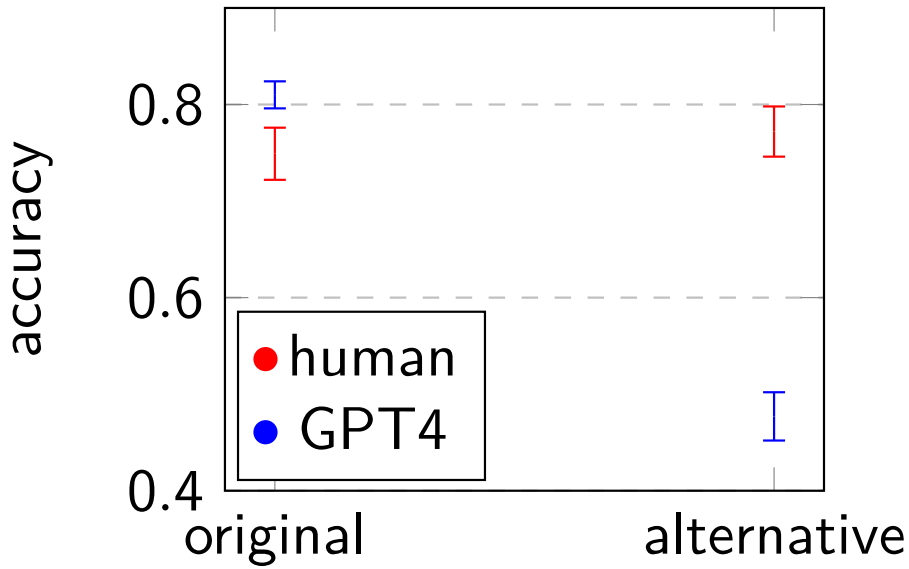


[28] Martha Lewis and Melanie Mitchell (2024). "Using Counterfactual Tasks to Evaluate the Generality of Analogical Reasoning in Large Language Models". In: *arXiv*

No, it doesn't

- different study on generalized analogical reasoning problems by Lewis and Mitchell[28]

- results for a number-matrix problem (results from Mitchell, 2024)



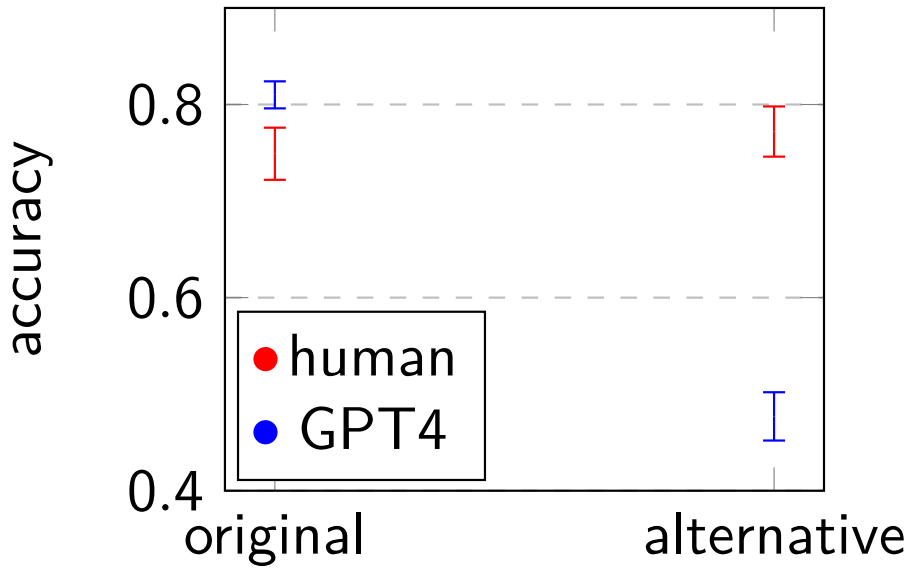
[28] Martha Lewis and Melanie Mitchell (2024). "Using Counterfactual Tasks to Evaluate the Generality of Analogical Reasoning in Large Language Models". In: *arXiv*

No, it doesn't

- different study on generalized analogical reasoning problems by Lewis and Mitchell[28]

→ being able to solve specific APs is no indicator for general intelligence

- results for a number-matrix problem (results from Mitchell, 2024)



[28] Martha Lewis and Melanie Mitchell (2024). "Using Counterfactual Tasks to Evaluate the Generality of Analogical Reasoning in Large Language Models". In: *arXiv*

Conclusion – from using APs in AI to using APs for AI

- in this talk focus on
  - solving APs with the help of AI and
  - understanding the underlying structure and formalization of APs
- different, even more important aspect: using AP for abstraction and generalization of AI-approaches to other topics, thus using AP *for* AI instead of *in* AI

Bibliography I


Antić, Christian (2022). "Analogical proportions". In: *Annals of Mathematics and Artificial Intelligence* 90.6, pp. 595–644.


Aristotle (1996). *Poetics*. Penguin Classics.


Bordes, Antoine, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko (2013). "Translating Embeddings for Modeling Multi-relational Data". In: *NIPS'13*, pp. 2787–2795.


Bounhas, Myriam and Henri Prade (2019). "An analogical interpolation method for enlarging a training dataset". In: *Scalable Uncertainty Management: 13th International Conference, SUM 2019, Compiègne, France, December 16–18, 2019, Proceedings 13*. Springer, pp. 136–152.


Correa, William Fernando, Henri Prade, and Gilles Richard (2012). "When intelligence is just a matter of copying". In: *European Conference on Artificial Intelligence*.


Dubois, Didier, Henri Prade, and Gilles Richard (2016). "Multiple-valued extensions of analogical proportions". In: *Fuzzy Sets and Systems* 292, pp. 193–202.


Falkenhainer, Brian, Kenneth D. Forbus, and Dedre Gentner (1989). "The structure-mapping engine: Algorithm and examples". In: *Artificial Intelligence* 41.1, pp. 1–63.


Forbus, Kenneth D. (2019). *Qualitative Representations: How People Reason and Learn about the Continuous World*. The MIT Press.


Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge (2016). "Image Style Transfer Using Convolutional Neural Networks". In: *2016 CVPR*, pp. 2414–2423.

Bibliography II

Hertzmann, Aaron, Charles E. Jacobs, Nuria Oliver, Brian Curless, and David H. Salesin (2001). "Image analogies". In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*.

Herzig, Andreas, Emiliano Lorini, and Henri Prade (2024). "A Novel View of Analogical Proportion Between Formulas". In: *ECAI 2024*.

Hodges, Wilfrid (2001). "Formal Features of Compositionality". In: *Journal of Logic, Language and Information* 10.1, pp. 7–28.

Hüllermeier, Eyke (2020). "Towards Analogy-Based Explanations in Machine Learning". In: *Modeling Decisions for Artificial Intelligence*, pp. 205–217.

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros (2017). "Image-to-Image Translation with Conditional Adversarial Networks". In: *2017 CVPR*, pp. 5967–5976.

Kant, Immanuel (2004). *Immanuel Kant: Prolegomena to Any Future Metaphysics*. Ed. by Gary Hatfield. 2nd ed. Cambridge University Press.

Leemhuis, Mena and Özgür L. Özçep (2023). "Analogical Proportions and Betweenness". In: *Proceedings of the 9th Workshop on Formal and Cognitive Reasoning, co-located with the 46th German Conference on Artificial Intelligence (KI 2023)*, pp. 8–19.

Leemhuis, Mena, Diedrich Wolter, and Özgür L. Özçep (2024). "Decomposing Analogy: A Logic Characterization". In: *Foundations of Information and Knowledge Systems*, pp. 256–274.

Lepage, Yves (2002). "Analogy and Formal Languages". In: *Electronic Notes in Theoretical Computer Science* 53, pp. 180–191.

Bibliography III


Lewis, Martha and Melanie Mitchell (2024). "Using Counterfactual Tasks to Evaluate the Generality of Analogical Reasoning in Large Language Models". In: *arXiv*.


Lieber, Jean, Emmanuel Nauer, Henri Prade, and Gilles Richard (2018). "Making the Best of Cases by Approximation, Interpolation and Extrapolation". In: *Case-Based Reasoning Research and Development*, pp. 580–596.


Lovett, Andrew M., Kenneth D. Forbus, and Jeffrey M. Usher (2010). "A Structure-Mapping Model of Raven's Progressive Matrices". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 32.


Małkiński, Mikolaj and Jacek Mańdziuk (2022). "Deep Learning Methods for Abstract Visual Reasoning: A Survey on Raven's Progressive Matrices". In: *ArXiv* abs/2201.12382.


Mitchell, Melanie (2024). *AI's Challenge of Understanding the World*. AI Spotlight Seminar. URL: https://www.youtube.com/live/gRJgqSbk71o.


Prade, Henri and Gilles Richard (2014). "From Analogical Proportion to Logical Proportions: A Survey". In: *Studies in Computational Intelligence*, pp. 217–244.


— (2018). "Analogical proportions: from equality to inequality". In: *International Journal of Approximate Reasoning* 101, pp. 234–254.


Raven, John C. (1941). "Standardzation of progressive matrices". In: *British Journal of Medical Psychology* 19.1, pp. 137–150.


Rumelhart, David E and Adele A Abrahamson (1973). "A model for analogical reasoning". In: *Cognitive Psychology* 5.1, pp. 1–28.

Bibliography IV

Ushio, Asahi, Luis Espinosa Anke, Steven Schockaert, and Jose Camacho-Collados (2021). "BERT is to NLP what AlexNet is to CV: Can Pre-Trained Language Models Identify Analogies?" In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, pp. 3609–3624.

Webb, Taylor, Keith J. Holyoak, and Hongjing Lu (2023). "Emergent analogical reasoning in large language models". In: *Nature Human Behaviour* 7.9, pp. 1526–1541.

Zhang, Chi, Feng Gao, Baoxiong Jia, Yixin Zhu, and Song-Chun Zhu (2019). "RAVEN: A Dataset for Relational and Analogical Visual REasoNing". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5312–5322.

Zhuo, Tao and Mohan S. Kankanhalli (2020). "Solving Raven's Progressive Matrices with Neural Networks". In: *ArXiv* abs/2002.01646.