AMSI SUMMERRESEARCH SCHOLARSHIPS 2023–24

SET YOUR SIGHTS ON RESEARCH THIS SUMMER

Diploma Seminar

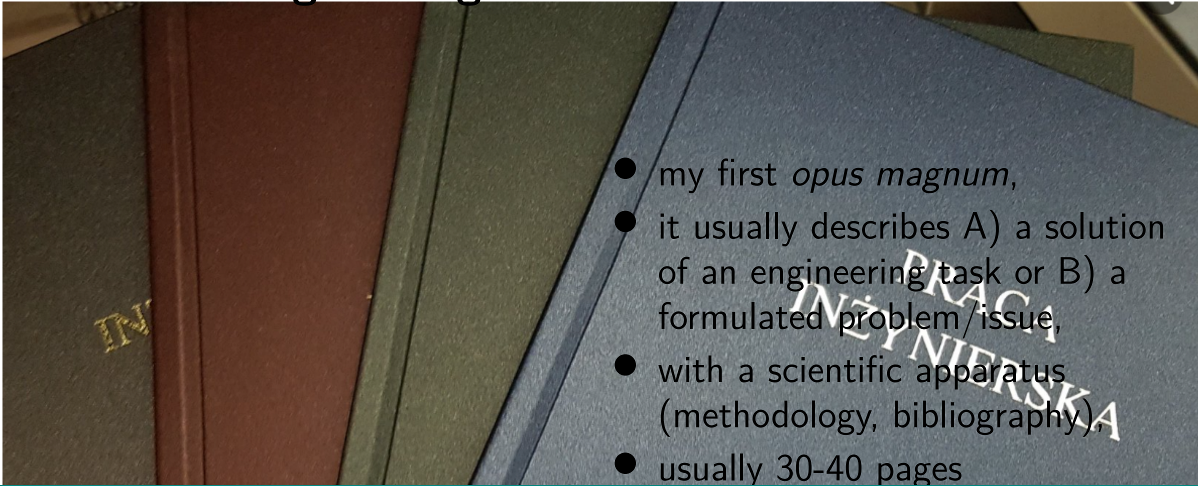
Krystian Jobczyk

AGH University

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

AMSI CONNECT

Engineering Thesis – what is about?



- my first *opus magnum*,
- it usually describes A) a solution of an engineering task or B) a formulated problem/issue,
- with a scientific apparatus (methodology, bibliography),
- usually 30-40 pages

- **A classical (ideal) structure of the thesis**
  https://missamandawatson.blogspot.com/2021/07/how-to-write-phd-thesis

How to begin??????

- Indicate which area of my study cycle is '**mine**'/**interesting**/**intriguing**/**useful**
- Find a potential supervisor (engaged/active in my favourite area) willing to supervise me
- Explore the background/recognise the subproblems/subtasks to be perform
- Make a literature overview + formulate the questions/problems

Some General, but (Potentially) Useful Hints:

- **Prepare your thesis achronologically:)**
  1. 'Core' of the thesis (experiments/data analysis)
  2. Theoretic background (dictionary, definitions, methods)
  3. Conclusions (+ open problems, closing remarks, further perspectives)
  4. Introduction (objective, structure, motivation)
- **A well-written thesis forms a compromise between:**
  1. conciseness and verboseness
  2. essay and encyklopedic compendium
  3. a popular science text and a 'jargon' expert one

Why I shouldn't start from **Introduction**? **You can**, ...but be aware that:

- **your analysis will run differently than you announce in 'Introduction'** – e.g. because:
  1. not all experiments appear noteworthy, fortunate, etc.– against your initial intuitions and goals
  2. you detect/discover something intriguing and initially unexpected, etc.

Why I shouldn't work on **Theoretic background** before proper analysis and experiments? **You can**, ...but be aware that:

- **it is difficult to initially indicate all the concepts and methods – needed in your experiments and analysis**

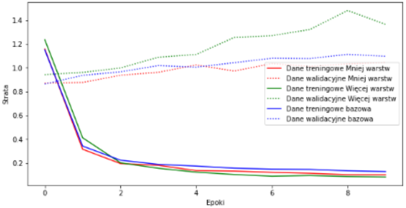
Core of the thesis: experiments/data analysis

- Be possibly **precise** and **informative** in your data/experiments presentation
- Take care of **interference** between your data **visualisation** (diagrams, pseudocodes, tables, etc.) and their **text description**.

Experiments/data analysis

Sieć	Ilość parametrów	Czas trenowania	Dokładność	Strata
Bazowa	5,176,458	143 s	0.7930	1.0876
Mniej neuronów	1,296,274	**34.5 s**	**0.7992**	1.0372
Więcej neuronów	10,562,370	360 s	0.7867	1.2342
Mniej warstw	5,144,110	142 s	0.7970	**0.9865**
Więcej warstw	5,282,186	110 s	0.7903	1.4197

Tabela 3.2. Porównanie wyników wg złożoności sieci



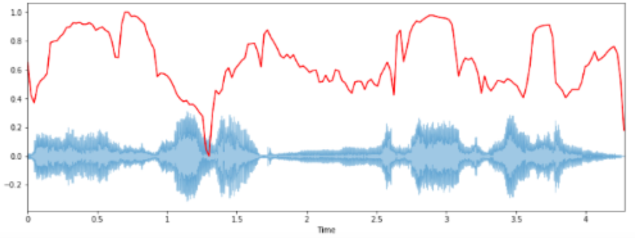
Rys. 3.7. Wykres straty dla różnych ilości warstw sieci. Źródło: Opracowanie własne

Evaluation

- Data – presented
- The diagram axes and units – named
- The role of colours –explained

Experiments/data analysis

Evaluation



Rys. 2.9. Wizualizacja spadku spektralnego utworu Krzysztof Krawczyk - Chciałem być

- Data – presented
- The diagram axes and units – (partially) named
- The colour-visualized data – only partially indicated

Experiments/data analysis

The model-checking for a formula ϕ in a multi-modal Kripke frame M, in its state ω, is given by Algorithm 1. It specifies a unique ε-dependent procedure VERIFY$(M, \omega, \phi, *, \varepsilon)$. Since the procedure is standard for atomic formulae, we omit its description in Algorithm 1. We only define the procedure for modal formula $\beta = \Diamond_i \gamma$, and $\beta = \langle BI \rangle^{\varepsilon} \gamma$, for some γ atomic.

In the case $\beta = \Diamond_i \gamma$ (lines 2-6 of Algorithm 1), the procedure works for every $\omega \in W$. If one can existentially choose υ such that $\omega <_i \upsilon$, for a given ω, and the procedure (positively) verifies that γ is satisfied in M, υ (we denote it by VERIFY$(M, \upsilon, \gamma, \varepsilon)$), then the procedure returns 'true'. Otherwise – it returns 'false'. In the case $\beta = \langle BI \rangle^{\varepsilon} \gamma$ (lines 7-14 of Algorithm 1), the procedure works for every $\omega \in M$, for arbitrarily large, but finite set of $n \in \{1, \dots, N\}$ and $\varepsilon \in \{a_1, \dots, a_l\}$, for some natural l. If we can find such k that every $n \leq k$ there exists such $u \in W$ that the

Algorithm 1 Procedure VERIFY$(M, \omega, \phi, *, \varepsilon)$, where
$* = \{<_i, BI_1^{<\varepsilon}, \dots, BI_n^{<\varepsilon}, Move_1^{<\varepsilon}, \dots, Move_n^{<\varepsilon}, Val)\}, i \in I, n \in \{1, \dots, N\}, \varepsilon \in \{a_1, \dots, a_l\}, \alpha$

1: **for all** $\beta \in Sub(\alpha)$ such that $|\beta| = i$ **do**
2: **while**
$\bullet \beta = \Diamond_i \gamma$
for every $\omega \in M, i \in I$ **do**
3: **if** $\exists \upsilon . \omega <_i \upsilon \wedge \gamma \in V(\upsilon)$ **then**
4: set := $Val(\omega) = Val(\omega) \cup \{\beta\}$ and VERIFY$(M, \upsilon, *, \gamma, \varepsilon)$
return true
5: **end if**
return false
6: **end while**
7: **while**
$\bullet \beta = \langle BI \rangle^{<\varepsilon} \gamma$
for every $\omega \in M, n \in \{1, \dots, N\}, \varepsilon \in \{a_1, \dots, a_l\}$ **do**
8: **if** $\exists k . \forall n \geq k\ \exists_{u \in W} (\omega BI^{<\varepsilon} u) \wedge \gamma \in Val(u)$ **then**
9: set := $Val(\omega) = Val(\omega) \cup \{\beta\}$ and VERIFY$(M, \omega, *, \gamma, \varepsilon)$ and
return true
10: **end if**
return false
11: **end while**
....
12: **end for**

Evaluation

- The Model-checking pseudocode presented
- A general description is given
- The key pseudocode line – explained in the text

Experiments/data analysis

Algorithm 1 Procedure VERIFY$(M, \omega, \phi, *, \varepsilon)$, where
$* = \{<_i, BI_1^{<\varepsilon}, \dots, BI_n^{<\varepsilon}, Move_1^{<\varepsilon}, \dots, Move_n^{<\varepsilon}, Val)\}, i \in I, n \in \{1, \dots, N\}, \varepsilon \in \{a_1, \dots, a_l\}, \alpha$

for all $\beta \in Sub(\alpha)$ such that $|\beta| = i$ **do**
  while
  $\bullet\beta = \Diamond_i \gamma$
  for every $\omega \in M, i \in I$ **do**
    if $\exists \upsilon . \omega <_i \upsilon \wedge \gamma \in V(\upsilon)$ **then**
      set := $Val(\omega) = Val(\omega) \cup \{\beta\}$ and VERIFY$(M, \upsilon, *, \gamma, \varepsilon)$
      return true
    end if
    return false
  end while
  while
  $\bullet\beta = \langle BI \rangle^{<\varepsilon} \gamma$
  for every $\omega \in M, n \in \{1, \dots, N\}, \varepsilon \in \{a_1, \dots, a_l\}$ **do**
    if $\exists k . \forall n \geq k \ \exists_{u \in W} \ (\omega BI^{<\varepsilon} u) \wedge \gamma \in Val(u)$ **then**
      set := $Val(\omega) = Val(\omega) \cup \{\beta\}$ and VERIFY$(M, \omega, *, \gamma, \varepsilon)$ and
      return true
    end if
    return false
  end while
  ....
end for

Evaluation

- The Model-checking pseudocode presented
- No general description is given
- No pseudocode line explained in the text

Why are these two aspects so important?

- Your analysis is **transparent** and **comprehensible**,
- Your experiments are **verifiable/falsifiable**,
- Your work gains a **didactic value** (a reader might repeat your experiments and might learn from you.)

Thanks to this, there is more **science** inside!

Theoretic Background-Definitions

Definition – a statement (usually expressed in a natural language) of the meaning of a term (a word, phrase, or other set of symbols).

Types of definitions:

1. intensional (explain the meaning of the statements)
   Ex: A *formal language* is a subset of a set of words over a given alphabet.
2. extensional (list the objects that the term describes)
   Ex: A set of PC formulae is a set given by the grammar: ϕ, $\neg\phi$, $\phi \wedge \psi$, $\phi \to \psi$, $\phi \to \psi$.

Due to another criterion:

Types of definitions:

1. classical (introduces a unique equality between a defined term and a phrase which explains the term)
   Ex: A *circle* is a subset of points on a plane equidistant from a given centre.
2. oxtensive (convey the meaning of a term by pointing out examples.)
   Ex: This is the R-code for linear regression for two given variables x and y:

```
lm(formula = y ~ x)
```

Theoretic Background-Definitions

Due to another criterion:

Types of definitions:

1. real (refers to the defined objects)
   Ex: A *circle* is a subset of points on a plane equidistant from a given centre.
2. nominal (refers to the names of the defined objects)
   Ex: A word "*circle*" means a subset of points on a plane equidistant from a given centre.
3. projective (project some concepts, establish some arrangements)
   Ex: By the phrase 'declarative languages' we mean here PROLOG and ASP.

Typical Errors in Defining:

1. *Idem per Idem:*

   Ex 1:

   "Logic is a research branch on reasoning due to logical rules".

   Ex 2:

   "Mathematics is a research branch on different mathematical objects".

2. *Ignotum per Ignotum:*

   Ex 3: "Conjunction is a unique two-argumental functor of PC"

Typical Errors in Defining:

1. *Ignotum per Ignotus:*

Neighborhood – - Set of nearby, usually adjacent cells, surrounding the considered cell. It is typically defined by relative positions respecting the position of the central cell. The neighborhood can be also defined in various ways and while in most cases it is defined by adjacency, it not necessarily needs to be close in Cartesian understanding. Using other words, neighborhood is the definition of what cells are used by a transition function when calculating the state in the proceeding iteration. The most common types of neighborhood used in the field of cellular automata are: *von Neumann neighborhood* and *Moore neighborhood*. The previous one consist of four cells that are orthogonally adjacent [2], the second, in addition, includes four, diagonally adjacent cells [3].

Ex 4:

Typical Errors in Defining:

1. *Categorical Translation:*

   Ex 5: "Mathematics is a class of all objects that are a subject of interest of mathematicians"

2. Impreciseness:

   Ex 6: A *convolution* is a mathematical operation on two functions (f and g) that produces a third function $f * g$ that expresses how the shape of one is modified by the other.

Typical Errors in Defining:

1. Impreciseness:

   Ex 6:"A *convolution* is a mathematical operation on two functions (f and g) that produces a third function $f * g$ that expresses how the shape of one is modified by the other".

2. Falsehood:

   Ex 7: "Declarative languages are the programming languages in which each object forms a class".

How to omit the errors?:

❶ **Imprecise definition:**

Ex 6:"A *convolution* is a mathematical operation on two functions (f and g) that produces a third function $f * g$ that expresses how the shape of one is modified by the other".

❷ **Precise definition:**

Ex 6*: Let f, g be Lebesque integrable function on $\mathbb{R}$ (symb. $f, g \in L^1(\mathbb{R}))^1$, and let $Con : L^1(\mathbb{R}) \times L^1(\mathbb{R}) \to L^1(\mathbb{R})$ be a mapping associating to the pair (f, g) a new function

$$f \star g(t) = \oint f(t)g(x - t)\mathrm{d}t \in [0, 1].$$

The function $f \star g(t)$ is said to be called *a convolution* of f and g.

How to omit the errors?:

❶ Precise definition:

Ex 6*:

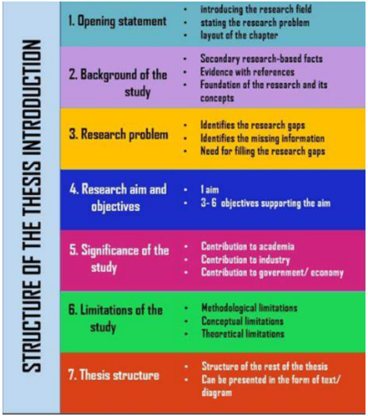
Definition 8: (**Convolution.**) Let f, g be Lebesgue integrable function on $\mathbb{R}$ (symb. $f, g \in L^1(\mathbb{R}))^6$, and let $Con : L^1(\mathbb{R}) \times L^1(\mathbb{R}) \to L^1(\mathbb{R})$ be a mapping associating to the pair (f, g) a new function

$$f * g(t) = \oint f(t)g(x-t)dt \in [0,1].$$

The function $f * g(t)$ is said to be called *a convolution* of f and g□

- **Hint**: Since your definitions will usually be technical concepts, they may require additional support definitions to be precise.

Introduction of the Thesis



- An exemplary structure of the '**Introduction**'

Introduction of the Thesis

1. **Motivation**:
   - It explains why the problem is formulated, why is interesting, noteworthy, etc.)
   - It usually infers the need to formulate the problem from the lacks and shortcomings of previous approaches.

In his groundbreaking article [1], Per Lindstrøm characterized the first-order logic as the maximal logic which satisfies both Compactness Theory and the Skolem-Löwenheim Theorem. These ideas initiated and influenced many other attempts at a similar characterizability of other classes of logical systems, such as modal logic systems in van Benthem's and de Rijke's approach in [2] and [3] (resp.) or the so-called infinitary chain logic systems as in Shelah's logic in [4] and – very recently – in Džamonja-Väänänen's approach [5].

Although the abstract model-theoretic treatment of analytic structures, such as Banach spaces, has been well-grounded since such works as [6], [7], their Lindstrøm's-type characterisation has still been elaborated for a small number of types of the structures, such as continuous metric spaces [9], [12] or – for structures with measures and integrals as in [11]. Unfortunately, no attempt to Lindstrøm's characterizability of structures of real and compact analysis has been yet proposed. In particular, we need more information about the appropriate maximal logic suitable to describe analytic

Introduction of the Thesis

1. **Objectives**:
   - They explain what is/are the thesis goal(s))
   - They usually refer to the motivation factors.

Ex:

Against this dominating trend in the research tradition and in some reference to the problems and questions, formulated above, the article is aimed at the following goals.

G1: To propose an (almost) fuzzy logic of action and preferences AFLAP – introduced both syntactically and semantically from its non-fuzzy 'surrogate' interpreted over finite game trees (breaking $\mathbf{Tend}_2$). (breaking $\mathbf{Tend}_3$).

G2: To prove the decidability of AFLAP using advanced and modified machinery of quasi-models (breaking $\mathbf{Tend}_1$).

G3: To give an outline of potential applications of AFLAP in 3 areas: a) in decision-making scenarios and b) in automated reasoning.

2.	**Czy treść pracy odpowiada tematowi określonemu w tytule?**
	b) Częściowo
3.	**Czy struktura pracy – podział treści, kolejność rozdziałów, itp. – jest odpowiednia?**
	a) Tak
4.	**Czy dobór źródeł oraz ich wykorzystanie są właściwe?**
	a) Tak
5.	**Czy praca jest poprawna językowo?**
	a) Tak
6.	**Czy technika edycji, spis rzeczy, odsyłacze – są poprawne?**
	a) Tak
7.	**Dodatkowe uwagi dotyczące pracy**
	W niektórych miejscach opisy są zbyt lapidarne, wręcz encyklopedyczne. Można by p ukonkretniające rozwinięcia.
	W przypadku braku dodatkowych uwag powyższe pole proszę uzupełnić słowem „brak".
8.	**Ocena pracy**
	4,5

- **How does your reviewer evaluate your thesis?**
- She/he has to answer to the following questions:

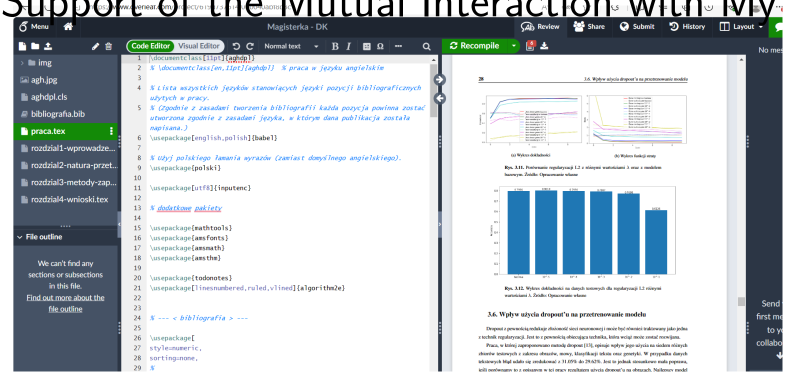
- Before that, she/he creates a review and indicates **positive** and **negative** sides of the thesis:

1.	**Merytoryczna ocena pracy**
	Praca dotyczy zagadnienia detekcji korelacji między typami utworów muzycznych a rodzajami emocji, jakie te utwory wzbudzają u słuchaczy. Tak określony cel pracy został w pełni zrealizowany poprzez konstrukcję modelu predykcyjnego, bazującego na sieciach neuronowych. Autor zastosował tu swoiste podejście porównawcze, wykorzystując dwa typy sieci neuronowych jako podstawę konstrukcji modelu. Do predykcji wykorzystał autor ponad 10 tysięcy ścieżek dzwiękowych z
	Nyquista, są uznane za powszechnie znane (chyba niesłusznie). Zaletą podejścia autora - oprócz pomysłu aproksymacji dobrych rezultatów poprzez użycie innego rodzaju sieci neuronowej - jest też pewna szczerość prezentacji. Autor podaje fragmenty kodu, np. dla reprezentacji utworu w bazie Google Colab. Zdaje się jednak, że brakuje pewnych elementów opisu, które pozwoliłyby czytelnikowi w szczegółach powtórzyć eksperymenty autora.

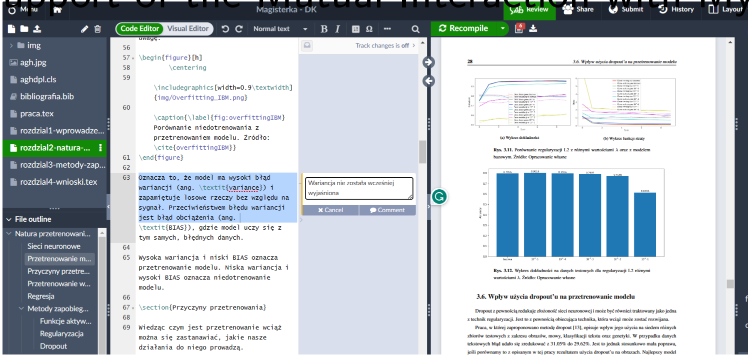
Technical Support of the Mutual Interaction with My Supervisor

Technical Support of the Mutual Interaction with My Supervisor

Technical Support of the Mutual Interaction with My Supervisor



Technical Support of the Mutual Interaction with My Supervisor



ENGINEERING EXAM

1. **Termin:** c.a. od 15/16 stycznia przez następne 2 tygodnie, sloty 9:00-12:00, 15:00-18:00
2. **Forma:** Ustna: 5-7 min prezentacji + 3 pytania Komisji + Ocena i Ogłoszenie wyników
3. 1 pytanie dotyczy pracy ("obronne"), 2 dotyczą programu studiów

EXAM-FINAL GRADES

Your final grade is a **weighted average** of the following components:

1. 60% – the average grade from the partial grades from the study period
2. 20% – the average grade from the diploma thesis evaluation
3. 20% – the average grade from the diploma exam

EXAM – FINAL GRADES

Oceny, o których mowa w rozdz. 3, a także wynik ukończenia studiów ustala się do dwóch miejsc po przecinku, bez zaokrągleń, zgodnie z następującą zasadą w zależności od wartości liczbowej:

1) od 3,00 ocena słowna: dostateczny (3.0)
2) od 3,21 ocena słowna: plus dostateczny (3.5)
3) od 3,71 ocena słowna: dobry (4.0)
4) od 4,21 ocena słowna: plus dobry (4.5)
5) od 4,71 ocena słowna: bardzo dobry (5.0).

Figure: The Method of Establishing the Final Exam Grades

EXAM – HOW TO ANSWER IN A CLEVER WAY

1. (The 'defence' question): Feel free to create your own narration, answer broadly, initiate some references, play with the associations, and indicate some interesting aspects of the problem.
2. (The 'exam' questions:):
   1. try to be precise and concise
   2. do not forget about examples (1-2)
   3. use an academic language (at least the typical vocabulary)
   4. try to not use any 'Polish-English" mix

1. try to be precise and concise....

= say (only) what is needed to answer your question/what correctly exposes the nature and properties of the objects you are being asked about.

Example

An Algorithm is a finite sequence of rigorous instructions, typically used to solve a class of specific problems or to perform a computation, satisfying the properties: A) correctness, B) finiteness, C) extensionality, D) effectiveness.

Example

I. An Algorithm is a finite sequence of rigorous instructions, typically used to solve a class of specific problems or to perform a computation, satisfying the properties: A) correctness, B) finiteness, C) extensionality, D) effectiveness.

II. We indicate the following algorithmic techniques (methods of their constructions:)

1. brute force, divide and conquer, greedy, heuristic, dynamic, evolutionary.

III. We can single out two important classes of algorithms:

1. searching algorithms and sorting algorithms

IV. Examples: Fermat's algorithm, Dijkstra's algorithm, Euclidean algorithm, etc.

Errors by Answering....

1. **Distinctive properties of our objects instead of their Definitions**

Example

A predicate calculus is based on quantifiers.

2. **Semantic fields of our concepts is too narrow.**

Example

A database is a collection of relational data.

Errors by Answering....

❶ **Operational/ostensive describing of our objects instead of their classical defining**

Example

I used a database in Chapter 5, on a page 23 of my thesis.

❷ **A definition of our concepts is partially false.**

Example

Algorithm is an infinite procedure.

Errors by Answering....

1. **Descriptions involved in classical defining errors (*idem per idem/ignotum per ignotum*)**

Example

A cell neighbourhood is a spatial structure consisting of a finite collection of orthogonally adjacent cells.

Errors by Answering....

1. **Defining involved in the error of Categorial Translation**

Example

A propositional calculus is a set of axioms.

2. **Defining by Applications**

Example

The relation might be found in a relational database.

What to do in difficult situations?

1. I didn't understand the question (or I am not sure that I understand it correctly)
2. I know nothing about the subject of the question/I completely forgot

AMSI SUMMER RESEARCH SCHOLARSHIPS 2023–24

SET YOUR SIGHTS ON
RESEARCH
THIS SUMMER

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

AMSI CONNECT