First Committee Meeting Progress Report

Jason Balaci

McMaster University

Oct. 21st, 2021

Table of Contents

- Introduction
- 2 Project
 - Drasil
 - Goal #1: Typed Expression Language
 - Goal #2: Model Discrimination "ModelKinds"
- References

Table of Contents

- Introduction
- 2 Project
 - Drasil
 - Goal #1: Typed Expression Language
 - Goal #2: Model Discrimination "ModelKinds"
- 3 References

• I am Jason Balaci



Me, Camping in Killarney Prov. Park, Fall 2019

- I am Jason Balaci
- Graduate of *McMaster University*, holding...



Me, Camping in Killarney Prov. Park, Fall 2019

- I am Jason Balaci
- Graduate of *McMaster University*, holding...
 - Hons. Actuarial and Financial Mathematics (B.Sc.)



Me, Camping in Killarney Prov. Park, Fall 2019

- I am Jason Balaci
- Graduate of McMaster University, holding...
 - Hons. Actuarial and Financial Mathematics (B.Sc.)
 - Minor in Computer Science



Me, Camping in Killarney Prov. Park, Fall 2019

- I am Jason Balaci
- Graduate of McMaster *University*, holding...
 - Hons. Actuarial and Financial Mathematics (B.Sc.)
 - Minor in Computer Science
- Currently pursuing a thesis-based Master's of Computer Science (M.Sc) at McMaster University, under the supervision of Dr. Jacques Carette.



Me, Camping in Killarney Prov. Park. Fall 2019

Course-related progression

• I'm required to complete¹²:

 $^{{\}it 1\atop https://academic calendars.romcmaster.ca/preview_program.php?catoid=45\&poid=23470\&returnto=9166}$

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

- I'm required to complete¹²:
 - One (1) "Software" course

 $^{{\}it 1\atop https://academic calendars.romcmaster.ca/preview_program.php?catoid=45\&poid=23470\&returnto=9166}$

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

- I'm required to complete¹²:
 - One (1) "Software" course
 - Fither of:

 $^{{\}it 1\atop https://academic calendars.romcmaster.ca/preview_program.php?catoid=45\&poid=23470\&returnto=9166}$

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

Course-related progression

- I'm required to complete¹²:
 - One (1) "Software" course
 - Either of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses

5/26

¹ https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

- I'm required to complete¹²:
 - One (1) "Software" course
 - Fither of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:

¹ https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

- I'm required to complete¹²:
 - One (1) "Software" course
 - Fither of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:
 - CAS 701 "Logic & Discrete Mathematics" Theory course, Fall 2020

¹ https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

- I'm required to complete¹²:
 - One (1) "Software" course
 - Fither of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:
 - CAS 701 "Logic & Discrete Mathematics" Theory course, Fall 2020
 - CAS 761 "Generative Programming" Software course, Fall 2020

l https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

² http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf

- I'm required to complete¹²:
 - One (1) "Software" course
 - Fither of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:
 - CAS 701 "Logic & Discrete Mathematics" Theory course, Fall 2020
 - CAS 761 "Generative Programming" Software course, Fall 2020
 - CAS 763 "Certified Programming with Dependent Types" Theory & Software course. Winter 2021

http://www.cas.mcmaster.ca/cas/Ofiles/reg_master_cs_2019a.pdf Jason Balaci (McMaster University) Committee Meeting 1



l https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

Course-related progression

- I'm required to complete¹²:
 - One (1) "Software" course
 - Either of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:
 - CAS 701 "Logic & Discrete Mathematics" Theory course, Fall 2020
 - CAS 761 "Generative Programming" Software course, Fall 2020
 - CAS 763 "Certified Programming with Dependent Types" Theory & Software course, Winter 2021
 - COMPSCI 6TB3 "Syntax-Based Tools and Compilers" Systems course, Winter 2021

5/26

¹ https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

- I'm required to complete¹²:
 - One (1) "Software" course
 - Either of:
 - Two "Theory" courses, and one "Systems" course
 - One "Theory" course, and two "Systems" courses
- I've completed:
 - CAS 701 "Logic & Discrete Mathematics" Theory course, Fall 2020
 - CAS 761 "Generative Programming" Software course, Fall 2020
 - CAS 763 "Certified Programming with Dependent Types" Theory & Software course, Winter 2021
 - COMPSCI 6TB3 "Syntax-Based Tools and Compilers" Systems course, Winter 2021
- Together, the courses completed satisfies the "Courses Requirement" as mentioned in the academic calendar¹ and the "Regulations for the Computer Science M.Sc. Program" document².

https://academiccalendars.romcmaster.ca/preview_program.php?catoid=45&poid=23470&returnto=9166

Thesis/research-related Progression

 Conducted "full-time" research for at least 1 full semester (Spring/Summer 2021), and "part-time" research during courses.

Thesis/research-related Progression

- Conducted "full-time" research for at least 1 full semester (Spring/Summer 2021), and "part-time" research during courses.
- Continuing to research "full-time".

Thesis/research-related Progression

- Conducted "full-time" research for at least 1 full semester (Spring/Summer 2021), and "part-time" research during courses.
- Continuing to research "full-time".
- Attended a thesis defence to learn about what to expect from a thesis defence (and learn about their research).

Thesis/research-related Progression

- Conducted "full-time" research for at least 1 full semester (Spring/Summer 2021), and "part-time" research during courses.
- Continuing to research "full-time".
- Attended a thesis defence to learn about what to expect from a thesis defence (and learn about their research).
- Supervisory committee is formed, and we are currently having our first supervisory committee meeting.
 - Supervisor: Dr. Jacques Carette
 - Dr. Spencer Smith
 - Dr. Wolfram Kahl

Table of Contents

- 1 Introduction
- 2 Project
 - Drasil
 - Goal #1: Typed Expression Language
 - Goal #2: Model Discrimination "ModelKinds"
- 3 References

What is Drasil?

Drasil...



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

 $^{^{1}}_{\rm https://jacquescarette.github.io/Drasil/}$

What is Drasil?

Drasil...

• is managed by Dr. Carette & Dr. Smith.



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

¹ https://jacquescarette.github.io/Drasil/

What is Drasil?

Drasil...

- is managed by Dr. Carette & Dr. Smith.
- originates from the work of Dan Szymczak.



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

¹ https://jacquescarette.github.io/Drasil/

What is Drasil?

- is managed by Dr. Carette & Dr. Smith.
- originates from the work of Dan Szymczak.
 - Originally focused on scientific software (*Literate Scientific* Software).



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

¹ https://jacquescarette.github.io/Drasil/

What is Drasil?

- is managed by Dr. Carette & Dr. Smith.
- originates from the work of Dan Szymczak.
 - Originally focused on scientific software (*Literate Scientific* Software).
 - Focus expanded...



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

¹ https://jacquescarette.github.io/Drasil/

What is Drasil?

- is managed by Dr. Carette & Dr. Smith.
- originates from the work of Dan Szymczak.
 - Originally focused on scientific software (Literate Scientific Software).
 - Focus expanded...
- tries to "Generate All The Things"...



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

https://jacquescarette.github.io/Drasil/

What is Drasil?

Drasil...

- is managed by Dr. Carette & Dr. Smith
- originates from the work of Dan Szymczak.
 - Originally focused on scientific software (Literate Scientific Software).
 - Focus expanded...
- tries to "Generate All The Things"...
 - with a focus on research software.



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]

8/26

https://jacquescarette.github.io/Drasil/

What is Drasil?

- is managed by Dr. Carette & Dr. Smith.
- originates from the work of Dan Szymczak.
 - Originally focused on scientific software (*Literate Scientific* Software).
 - Focus expanded...
- tries to "Generate All The Things"...
 - with a focus on research software.
- has a website¹!



Drasil's Logo [Carette et al., 2021][Yggdrasil - Wikipedia, 2021]



Drasil

"Generate All The Things!"

• TODO: here!

Drasil Case Studies

 $^{^{1} \}mathtt{https://jacquescarette.github.io/Drasil/\#Sec:Examples}$

Drasil Case Studies

• Drasil currently contains a significant amount of Physics-related knowledge.

¹ https://jacquescarette.github.io/Drasil/#Sec:Examples

Drasil Case Studies

- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:

¹ https://jacquescarette.github.io/Drasil/#Sec:Examples

- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.



 $¹_{ t https://iacquescarette.github.io/Drasil/#Sec:Examples}$

- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
 - Single Pendulum Observing the motion of a single pendulum.



- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
 - **Single Pendulum** Observing the motion of a single pendulum.
 - Double Pendulum Observing the motion of a double pendulum.



- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
 - **Single Pendulum** Observing the motion of a single pendulum.
 - Double Pendulum Observing the motion of a double pendulum.
 - Game Physics Modelling of an open source 2D rigid body physics library used for games.



- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
 - Single Pendulum Observing the motion of a single pendulum.
 - Double Pendulum Observing the motion of a double pendulum.
 - Game Physics Modelling of an open source 2D rigid body physics library used for games.
 - Proportional Derivative Controller (PDController) Examining the output of a "Power Plant" (Process Variable) over time.



10 / 26

https://jacquescarette.github.io/Drasil/#Sec:Examples

- Drasil currently contains a significant amount of Physics-related knowledge.
- As of writing, current case studies¹ are primarily related to physics, including:
 - GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
 - Single Pendulum Observing the motion of a single pendulum.
 - Double Pendulum Observing the motion of a double pendulum.
 - **Game Physics** Modelling of an open source 2D rigid body physics library used for games.
 - Proportional Derivative Controller (PDController) Examining the output of a "Power Plant" (Process Variable) over time.
 - Solar Water Heating System (SWHS) Modelling of a solar water heating system with phase change material, predicting temperatures and change in heat energy of water and the PCM over time.

Committee Meeting 1

¹ https://jacquescarette.github.io/Drasil/#Sec:Examples ← □ ト ← 壹 ト ← 壹 ト → 壹 → → 壹 → → ⊙ へ ○

- cont d^1 .
 - SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.



 $^{^{1}}_{\tt https://jacquescarette.github.io/Drasil/\#Sec:Examples}$

- cont.d¹:
 - SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.
 - Projectile Determining if a launched projectile hits a target, assuming no flight collisions.



https://jacquescarette.github.io/Drasil/#Sec:Examples

- cont.d¹:
 - SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.
 - Projectile Determining if a launched projectile hits a target, assuming no flight collisions.
 - Slope Stability Analysis Program (SSP) Assessment of the safety
 of a slope (composed of rock and soil) subject to gravity, identifying
 the surface most likely to experience slip and an index of its relative
 stability (factor of safety).

• *cont.d*¹:

- SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.
- Projectile Determining if a launched projectile hits a target, assuming no flight collisions.
- Slope Stability Analysis Program (SSP) Assessment of the safety
 of a slope (composed of rock and soil) subject to gravity, identifying
 the surface most likely to experience slip and an index of its relative
 stability (factor of safety).
- Heat Transfer Coefficients between Fuel and Cladding in Fuel Rods (HGHC) - Examining the heat transfer coefficients related to clad.



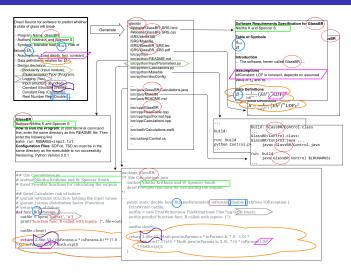
¹ https://jacquescarette.github.io/Drasil/#Sec:Examples

cont.d¹:

- SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.
- Projectile Determining if a launched projectile hits a target, assuming no flight collisions.
- Slope Stability Analysis Program (SSP) Assessment of the safety
 of a slope (composed of rock and soil) subject to gravity, identifying
 the surface most likely to experience slip and an index of its relative
 stability (factor of safety).
- Heat Transfer Coefficients between Fuel and Cladding in Fuel Rods (HGHC) - Examining the heat transfer coefficients related to clad.

The Drasil website is also generated by Drasil!

Taking a closer look at one of the examples: GlassBR



Knowledge flow from "knowledge-base"/source to artifacts, by Dr. Spencer Smith

GlassBR Generates Code!

 GlassBR - Predicting whether or not a glass slab is likely to resist a specified blast.

- GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
- Proportional Derivative Controller (PDController) Examining the output of a "Power Plant" (Process Variable) over time.

- GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
- Proportional Derivative Controller (PDController) Examining the output of a "Power Plant" (Process Variable) over time.
- SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.

- GlassBR Predicting whether or not a glass slab is likely to resist a specified blast.
- Proportional Derivative Controller (PDController) Examining the output of a "Power Plant" (Process Variable) over time.
- SWHS without Phase Change Material (NoPCM) Modelling of a solar water heating system without phase change material, predicting temperatures and change in heat energy of water and the PCM over time.
- **Projectile** Determining if a launched projectile hits a target, assuming no flight collisions.

Where will I be contributing?

Terminology is currently being changed, but is not reflected in many documents yet. () +

Where will I be contributing?

After all,

Jason Balaci (McMaster University)

Terminology is currently being changed, but is not reflected in many documents yet.

Where will I be contributing?

After all,

• They're all covered under "well-understood" domains!

Jason Balaci (McMaster University)

Terminology is currently being changed, but is not reflected in many documents yet:

Where will I be contributing?

After all,

- They're all covered under "well-understood" domains!
- The SRS documents are generated!

Where will I be contributing?

After all.

- They're all covered under "well-understood" domains!
- The SRS documents are generated!

Generating view-only data (e.g., SRS documents) is considerably easier than generating working code.

Where will I be contributing?

After all.

- They're all covered under "well-understood" domains!
- The SRS documents are generated!

Generating view-only data (e.g., SRS documents) is considerably easier than generating working code.

A few, notable, blocking problems:

Where will I be contributing?

After all.

- They're all covered under "well-understood" domains!
- The SRS documents are generated!

Generating view-only data (e.g., SRS documents) is considerably easier than generating working code.

A few, notable, blocking problems:

 Confidently generating usable software artifacts without strong type information places significant stress on developers, resulting in a higher likelihood of bugs in artifacts.

Where will I be contributing?

After all.

- They're all covered under "well-understood" domains!
- The SRS documents are generated!

Generating view-only data (e.g., SRS documents) is considerably easier than generating working code.

A few, notable, blocking problems:

- Confidently generating usable software artifacts without strong type information places significant stress on developers, resulting in a higher likelihood of bugs in artifacts.
- Existing "theories"/"*Models" don't expose enough information. They must be enriched, so that we can pull more information from them in straightforward manner.

The Problem

Goal #1: Typed Expression Language The Problem

• Ensure only admissible expressions are used in GOOL-supported languages, and that all expressions are coherent.

Goal #1: Typed Expression Language The Problem

- Ensure only admissible expressions are used in GOOL-supported languages, and that all expressions are coherent.
- Eases developer cognitive load when writing expressions, as they will need to ensure their expressions are coherent, or else a type error will be thrown

What makes up a "good" solution?

What makes up a "good" solution?

 Catches, within reason, all possible scenarios where an expression goes awry.

What makes up a "good" solution?

- Catches, within reason, all possible scenarios where an expression goes awry.
- Allows GOOL code generator to also become typed!

Current Progression

 Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)

- Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)
 - 'Expr' is a discrete, directly computable language (calculator possible)

- Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)
 - 'Expr' is a discrete, directly computable language (calculator possible)
 - Created 'ModelExpr' contains all possible expressions we would want to write on pencil-and-paper. There are still a few operations left in "Expr" that need to be moved over, however.

- Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)
 - 'Expr' is a discrete, directly computable language (calculator possible)
 - Created 'ModelExpr' contains all possible expressions we would want to write on pencil-and-paper. There are still a few operations left in "Expr" that need to be moved over, however.
 - Theories that rely on discussion of terms only found 'ModelExpr' may only have code generation made possible if we have rich enough data (see goal #2)

- Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)
 - 'Expr' is a discrete, directly computable language (calculator possible)
 - Created 'ModelExpr' contains all possible expressions we would want to write on pencil-and-paper. There are still a few operations left in "Expr" that need to be moved over, however.
 - Theories that rely on discussion of terms only found 'ModelExpr' may only have code generation made possible if we have rich enough data (see goal #2)
 - 'CodeExpr' contains a total conversion from 'Expr' with a few extra functionalities for GOOL

Current Progression

- Split core mathematical expression language (Expr) into 3 variants (Expr, ModelExpr, and CodeExpr)
 - 'Expr' is a discrete, directly computable language (calculator possible)
 - Created 'ModelExpr' contains all possible expressions we would want to write on pencil-and-paper. There are still a few operations left in "Expr" that need to be moved over, however.
 - Theories that rely on discussion of terms only found 'ModelExpr' may only have code generation made possible if we have rich enough data (see goal #2)
 - 'CodeExpr' contains a total conversion from 'Expr' with a few extra functionalities for GOOL
 - Created a "typed tagless final" smart constructor encoding for writing expressions in "Expr" (or, optionally, ModelExpr).

What are the next steps?

What are the next steps?

Continuing to remove terms.

What are the next steps?

- Continuing to remove terms.
- Moving literals from "Expr" & "ModelExpr" into their own small language, so that areas that want *strictly* literals can also have stronger restrictions on allowed data (terms).

Goal #2: Model Discrimination – "ModelKinds" The Problem

Goal #2: Model Discrimination – "ModelKinds" The Problem

• "RelationConcept"s don't contain enough information on their own to be a core component usable in code generation.

- "RelationConcept"s don't contain enough information on their own to be a core component usable in code generation.
- They were essentially "Relation"s ("Expr"s) with a natural language description of them.

- "RelationConcept"s don't contain enough information on their own to be a core component usable in code generation.
- They were essentially "Relation"s ("Expr"s) with a natural language description of them.
- If the "shape" of the expressions are not uniform, then writing more "interpreters"/"views"/code generators for them required difficult pattern analysis. It's also not a total-conversion.

What makes up a "good" solution?

What makes up a "good" solution?

 A good solution involves making the "Relation"s a "view" of a more data-rich specialized container for each kind of "*Model".

What makes up a "good" solution?

- A good solution involves making the "Relation"s a "view" of a more data-rich specialized container for each kind of "*Model".
- "ModelKinds"

What makes up a "good" solution?

- A good solution involves making the "Relation"s a "view" of a more data-rich specialized container for each kind of "*Model".
- "ModelKinds"
- By constructing our final data views through "more steps" (e.g., with more depth), we create a system of specialization, and we can start to see how specialization will occur.

Current Progression

Current Progression

 Considerable number of "theories"/"*Models" have been restructured, but there are still many that are pending typing. Most are best to be done once we have a typed expression language, and the rest are differential equation-related models (primarily Dong's domain).

What are the next steps?

What are the next steps?

 Understanding what kinds of needs we have for "collections", pushing this information back into the typed expression language (once that is fully typed), and then further creating model containers for these models.

What are the next steps?

- Understanding what kinds of needs we have for "collections", pushing this information back into the typed expression language (once that is fully typed), and then further creating model containers for these models.
- For the differential equation-related models, we will need to build appropriate models for each kind.

Acknowledgements

Fin.
Thank you!

Table of Contents

- Introduction
- 2 Project
 - Drasil
 - Goal #1: Typed Expression Language
 - Goal #2: Model Discrimination "ModelKinds"
- References

References I



Carette, J., Smith, S., Balaci, J., Hunt, A., Wu, T.-Y., Crawford, S., Chen, D., Szymczak, D., MacLachlan, B., Scime, D., and Niazi, M. (2021).

Drasil.



Yggdrasil - Wikipedia (2021).

Yggdrasil.