All Spreadsheets must Die

Robert Schadek May 8, 2019

Getting started

A random list of languages we love to hate



C++ 4.4*Million* (2015)
C 1.9*Million* (2015)
Java 9*Million* (2009)
JS 10*Million* (2018)

These are all small fish

[3, 4]

These are all small fish

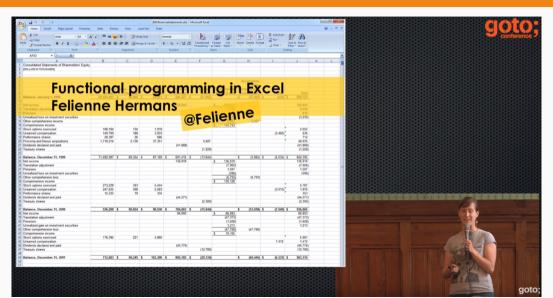
Excel

pprox 750 Million (2016)





Oh, but it is



A little bit of Spreadsheet bashing

B1 💌 🎉 ∑ 🚍 🖂 A1 * 13.37				
	A	В	С	
1	10	133.7		
2	14	187.18		
3				

	A	B	С	
1	Firstname	Lastname	Age	
2	John	Doe	34	
3	Hans	Mustermann	twenty-five	
4			-	

D4		- 🖌 Σ	💌 🚘 Σ 🚍		
	A	В	C		
1	Firstname	Lastname	Age		
2	John	Doe	34		
3	Hans	Mustermann	25		

СЗ		- 🖌 Σ	= 25	_
	A	В	С	
1	Firstname	Lastname	Age	
2	John	Doe		34
3	Hans	Mustermann		25



git blame

git blame

git blame

git blame

lets not go there

we will just become sad

=SUM(1,2)

- =SUM(1,2)
- equal, identifier, lparen, int(1), comma, int(2), rparen

- =SUM(1,2)
- equal, identifier, lparen, int(1), comma, int(2), rparen

set Excel locale to de_DE

- =SUM(1,2)
- equal, identifier, lparen, int(1), comma, int(2), rparen

- set Excel locale to de_DE
- =SUM(1,2)

- =SUM(1,2)
- equal, identifier, lparen, int(1), comma, int(2), rparen

- set Excel locale to de_DE
- =SUM(1,2)
- equal, identifier, lparen, float(1.2), rparen

- Knowledge silos
- Slow
- No separation between data and code
- Access management ...

- Knowledge silos
- Slow
- No separation between data and code
- Access management ... anybody?

- 1. Create private shopping spreadsheet
- 2. Show spreadsheet to college
- 3. Use spreadsheet for all company purchases
- 4. Put web frontend on spreadsheet backend
- 5. Pivot company to become E-Commerce company

- 1. Create private shopping spreadsheet
- 2. Show spreadsheet to college
- 3. Use spreadsheet for all com/ chases
- 4. Put web frontend on spread
- Excel snafu costs firm \$24m Some cleric, some error 5. Pivot company to become E-Co



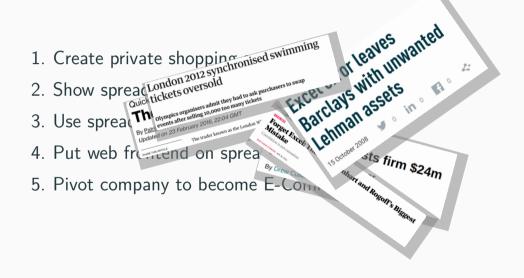


- Excel error leaves with unwanted Barclays with unwanted 1. Create private shopping spread- eet
- 3. Use sprea The London Whale
 - Updated on 23 February
- 4. Put web freed on sprea
- 5. Pivot company to become E-Con

15 October 2008

sts firm \$24m

And ROSOIT'S BIRSEST





Fannie Mae \$1.2bn honest mistake wanted RedEnvelope skids on loss forecast 1. Create private shopping und swimming 2: 2. Show spreacLondon 201 Join o Ho Barclay asses EI 💙 🖾 🔽 🛥 🗭 0 imated on budgeting error: CFO resigns 3. Use spread Updated on . 15 October 2008 4. Put web freed on sprea sts firm \$24m 5. Pivot company to become E-Com ^{Cand} Rogolf's Biggest





1. Create pr

2. Show sp

3. Use spr

4. Put web

5. Pivot cor

A format A form

SPRIG ORIGINAL HORROR STOR

The stories are listed in reverse chronological order.

389) Which way around do you calculate a percentage change?

http://www.startribune.com/587/story/367953.html

Making a political plus of math error, Star Tribune, 12 Apr 06

DFL er feleteca. Oto is accurate Justico PAR Anderson, a Republican, ot sloppy work. In a column reporting the precentage changes in unreserved that balances from 2020 to 2041, instead of duriding the difference is the 2020 Bgue. Republic Address Tayly Station said "The researcher who worked on that report just made a mistake in the formula in the separathenet. It heals bala about 8."

Risk: Giving political opponents an opportunity to comment adversely on the state auditor Avoidance: Check not just the data but the formulas too. Advice: Realister Novi for the EuSoRiO 2012 Conference

088) Scoring an own goal

http://www.channelregister.co.uk/2006/03/10/ogc_spreadsheet_snafu/

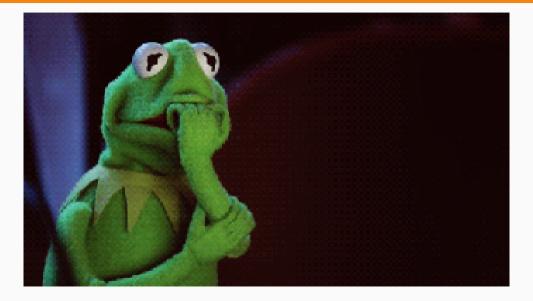
The Register (UK) OGC spreadsheet madness 10 Mar 06

The Office of Government Commerce is biaming a spreadsheet error for a foul-up over accrediting suppliers for its new Catalist procurement programme. In a keter to suppliers, they confessed "Unfortunately, live are] not, as we had hoped, in a position to acceet voor terefore at this time. This is because an error in the original evaluations secondshee has been identified. Increasifiant



Spreadsheets rule the world!

How you should be feeling right now



You believe that spreadsheets rule the world. You want D to rule the world instead.

How are we going to win this?

How are we going to win this?

We are not!

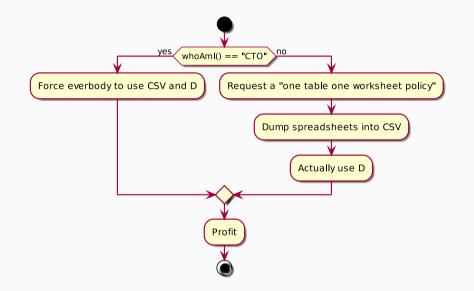
Lets draw up a battle plan

- Too many spreadsheets
- Too many tasks
- Too little man-power

- Too many spreadsheets
- Too many tasks
- Too little man-power

- Millions of lines of source in different languages
- D

Possible Attack Vectors



How to work with limited man-power

Leveraging existing libraries

- It is required, people will ask for that
- Writing a somewhat feature complete xlsx writer is a huge task

- It is required, people will ask for that
- Writing a somewhat feature complete xlsx writer is a huge task

- libxlsxwriter is a feature rich xlsx writer
- Wrapping it by hand,

- It is required, people will ask for that
- Writing a somewhat feature complete xlsx writer is a huge task

- libxlsxwriter is a feature rich xlsx writer
- Wrapping it by hand, no way (+78000 lines of structs, enums and functions)

- It is required, people will ask for that
- Writing a somewhat feature complete xlsx writer is a huge task

- libxlsxwriter is a feature rich xlsx writer
- Wrapping it by hand, no way (+78000 lines of structs, enums and functions)

- dpp to the rescue
- libxlsxwriter.d (+4000 lines)
- But it is still a C api

void chart_axis_set_name(lxw_chart_axis*, const(char)*)
void chart_axis_set_name_font(lxw_chart_axis*, lxw_chart_font*)
void chart_axis_set_num_font(lxw_chart_axis*, lxw_chart_font*)
void chart_axis_set_num_format(lxw_chart_axis*, const(char)*)
void chart_axis_set_line(lxw_chart_axis*, lxw_chart_line*)
void chart_axis_set_fill(lxw_chart_axis*, lxw_chart_fill*)
void chart_axis_set_fill(lxw_chart_axis*, lxw_chart_fill*)
void chart_axis_set_fill(lxw_chart_axis*, lxw_chart_fill*)

Semi-automatic refactoring

```
struct ChartAxis {
1
     lxw_chart_axis* handle;
2
3
     void setName(string name) {
4
       chart_axis_set_name(this.handle, toStringz(name));
5
     }
6
7
     void setNameRange(string n, lxw row t row,
8
         lxw_col_t col)
9
     Ł
10
       chart axis set name(this.handle, toStringz(n), row, col);
11
     }
12
13
     . . .
14 }
```

Problem to solve: We needed fake data with a variety of attributes.

- Name
- Address
- i18n
- • • •



faker.js

- +160 attributes
- 39 languages

```
module["exports"] = [
    "#{prefix} #{first_name} #{last_name}",
    "#{first_name} #{nobility_title_prefix} #{last_name}",
    "#{first_name} #{first_name} #{fir
```

Listing 1: locales/de/name/name.js

FakeD

```
override string nameName() {
1
     switch(uniform(0, 6, this.rnd)) {
2
       case 0:
3
         return format!"%s %s %s"(namePrefix(), nameFirstName().
4
              nameLastName());
5
       case 1:
6
         return format!"%s %s %s"(nameFirstName(), nameNobilityTitlePrefix().
7
             nameLastName()):
8
9
       case 2:
         return format!"%s %s"(nameFirstName(), nameLastName());
10
       case 3:
11
         return format!"%s %s"(nameFirstName(), nameLastName()):
12
       case 4:
13
         return format!"%s %s"(nameFirstName(), nameLastName());
14
15
       case 5:
         return format!"%s %s"(nameFirstName(), nameLastName());
16
       default: assert(false);
17
     }
18
```

```
28
```

[7]

```
import faked;
auto f = new Faker(1337);
writeln(f.nameName());
// localized to german
f = new Faker_de(1338);
writeln(f.nameName());
```



- Input:
 - Parser and Generator ≈ 1500 lines of D
 - A day of boring work



- Input:
 - Parser and Generator ≈ 1500 lines of D
 - A day of boring work
- Output:
 - Output feature equivalent \approx 70000 lines faker.js clone
 - Most changes in faker.js just require a rerun of the tool to update



- Input:
 - Parser and Generator ≈ 1500 lines of D
 - A day of boring work
- Output:
 - Output feature equivalent \approx 70000 lines faker.js clone
 - Most changes in faker.js just require a rerun of the tool to update
- Bonus:
 - Created two PRs to faker.js fixing wrong template expansion

Taking a step back

Firstname	Lastname	Amount	Currency	CreatedBy
Hans	Meier	73331	USD	Ruth Ember
John	Doe	83431	GPB	Ruth Ember
Ruth	Ember	103431	EUR	Hans Meier

The starting point

```
class Employee {
1
     long id;
2
     DateTime createdAt;
3
4
     EmployeeInfo info;
5
     long infoId;
6
7
8
   class EmployeeInfo {
9
10
     long id;
     string firstname;
11
     string lastname;
12
13
     Salary salary;
14
     long salaryId;
15
   }
16
```

17	class Salary {
18	long id;
19	Employee createdBy;
20	<pre>long createdById;</pre>
21	
22	CurrencyAmount amount;
23	long amountId;
24	}
25	
26	<pre>class CurrencyAmount {</pre>
27	long id;
28	double amount;
29	
30	Currency currency;
31	<pre>long currencyId;</pre>
32	}

32 class Currency {
33 long id;
34 string name;
35 }

- interface Backend {
- 2 Employee[] getAllEmployees();
- 3 Employee getEmployee(long empId);
- 4 EmployeeInfo getEmployeeInfo(long empInfoId);
- 5 Salary getSalary(long salaryId);
- 6 CurrencyAmount getCurrencyAmount(long amountId);
- 7 Currency getCurrency(long currencyId);
- 8 }

The Frontend Code: Types

```
interface Employee {
1
     id: number;
2
     createdAt: number;
3
4
     info?: EmployeeInfo;
5
     infoId: number;
6
  }
7
8
   interface EmployeeInfo {
9
     id: number:
10
    firstname: string;
11
12
    lastname: string;
13
     salary?: Salary;
14
     salaryId: number;
15
16 }
```

```
17 interface Salary {
```

```
18 id: number;
```

```
19 createdBy?: Employee;
```

```
20 createdById: number;
```

```
22 amount?: CurrencyAmount;
23 amountId: number;
```

```
24 }
```

25

29

30

31

32

}

21

```
26 interface CurrencyAmount {
27 id: number;
```

```
amount: number;
```

```
currency?: Currency;
```

```
currencyId: number;
```

```
import {
1
       Employee, EmployeeInfo, Salary, CurrencyAmount, Currency
2
  } from "model";
3
4
  class Backend {
5
       getAllEmployees() : Employee[] { ... }
6
       getEmployee(empId: number): Employee { ... }
7
       getEmployeeInfo(empInfoId: number): EmployeeInfo { ... }
8
       getSalary(salaryId: number): Salary { ... }
9
       getCurrencyAmount(amountId: number): CurrencyAmount { ... }
10
       getCurrency(currencyId: number): Currency { ... }
11
12 }
```

35

```
this.backend.getAllEmployees().pipe(
1
     mergeMap((emps: Employee[]) => {
2
       const obs = [];
3
       for(const emp of emps) {
4
         obs.push(this.backend.getEmployeeInfo(emp.infoId)
5
            .pipe(map((empInfo: EmployeeInfo) => {
6
                const ne: Employee = {...emp, info : empInfo};
7
                return ne:
8
              })
9
10
         );
11
       }
12
       return forkJoin(obs):
13
     }).
14
     mergeMap((emps: Employee[]) => {
15
       const obs = [];
16
```

The Communication

Frontend	Backend			
getA ll Employees()				
Employee[] allEmps				
foreach getEmployeeInfo(emp.infold);				
EmployeeInfo info				
foreach getSalary(info.salaryId);				
Salary sal				
foreach getCurrencyAmount(sal.amountld);				
CurrencyAmount ca				
foreach getCurrency(ca.currencyld);				
Currency c				
foreach getEmployee(sal.createdByld);				
Employee createdby				
foreach getEmployeeInfo(createdBy.infold);				
EmployeeInfo createdbyInfo				
Frontend	Backend			

- Clearly, this is unworkable
- Not plastic at all
- Just a lot of boring work

The real Takeway



Declare how we want data to get, when we're asking for it.

Why can't we write this?

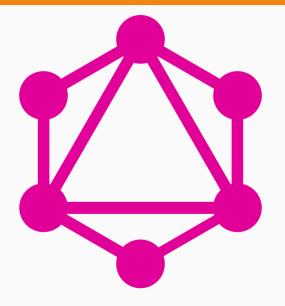
ſ allEmployees { info { firstname lastname salary { amount currency { name } createdBy { info { firstname lastname }

Why can't we write this?

1	{	1 {
2	allEmployees {	2 allEmployees: [{
з	info {	3 info: {
4	firstname	4 firstname: "Hans",
5	lastname	5 lastname: "Meier",
6	salary {	6 salary: {
7	amount	7 amount: 73331,
8	currency {	8 currency: {
9	name	9 name: "USD"
10	}	10 }
11	createdBy {	11 createdBy: {
12	info {	12 info: {
13	firstname	13 firstname: "Ruth",
14	lastname	14 lastname: "Ember"
15	}	15 }
16	}	16 }
17	}	17 }
18	}	18 },
19	}	19 },
20	}	20
		01]

41

GraphQL



GraphQL

```
1 schema {
     query: Backend
2
3 }
4
  type Backend {
5
     getAllEmployees: [Employee] 20
6
7
 - }
8
  type Employee {
9
     id: number!;
0
     createdAt: number!;
1
2
     info: EmployeeInfo;
3
     infoId: number!:
4
5 }
```

```
type EmployeeInfo {
15
      id: number!;
16
      firstname: String!;
17
      lastname: String!;
18
      salaryId: number!;
19
      salary: Salary;
21
  }
22
   type Salary {
23
      id: number!;
24
      createdBy: Employee;
25
      amountId: number!:
26
      amount: CurrencyAmount;
27
  }
28
```

```
type CurrencyAmount {
27
     id: number!;
28
     amount: number!:
29
      currencyId: number!;
30
      currency: Currency;
31
   }
32
33
   type Currency {
     id: number!;
34
     name: string!;
35
36
  }
```

Less code is better

```
1 query one {
     allEmployees {
2
        ...deep
3
     }
4
5
6
   fragment names on EmployeeInfo {
7
     firstname
8
     lastname
9
  3
10
11
   fragment empInfo on Employee {
12
     info {
13
     ...names
14
    <u>}</u>
15
   }
16
```

```
fragment deep on Employee {
1
      info {
2
        ...names
3
        salary {
4
          amount
5
          currency {
6
             name
7
          3
8
          createdBy {
9
             ...empInfo
10
          }
11
12
     }
13
14 }
```

Introspecting Types

```
ſ
1
       __type(name: "Employee") {
 2
         name
         fields {
 4
           name
 5
           type {
 6
 7
              name
              kind
 8
              ofType {
9
                 name
10
11
12
         }
13
14
    }
15
```

```
ſ
  "data": {
    "__type": {
      "name": "Employee",
      "fields": [
          "name": "id",
          "type": {
            "name": null.
             "kind": "NON_NULL",
             "ofType" {
               "name": "INT"
        },
          "name": "info",
          "type": {
             "name": "EmployeeInfo",
             "kind": "OBJECT"
```

-

45

```
20
```

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15 16 17

18

19



JUST DO IT



```
import graphqld;
1
2
   interface Query {
3
     Employee[] getAllEmployees();
4
   }
5
6
  class Schema {
7
     Query queryType;
8
  }
9
```

```
import graphqld;
1
2
  interface Query {
3
     Employee[] getAllEmployees();
4
   3
5
6
  class Schema {
7
     Query queryType;
8
  }
9
```

```
auto graphqld = new GraphQLD!(Schema)();
1
2
   graphqld.setResolver(
3
     "queryType", "getAllEmployees",
4
     delegate(string name, Json parent,
5
         Json args, ref Context context) @safe
6
     ſ
7
       Employee[] employees = getAllEmployees();
8
       Json ret = Json.emptyObject();
9
10
       ret["data"] = toGraphqlJson(employees);
11
       return ret;
12
     });
```

GraphQLD

```
void graphqlEndpoint(HTTPServerRequest req,
1
       HTTPServerResponse res)
2
  {
3
     string toParse = extractQuery(req);
4
     auto p = Parser(Lexer(toParse));
5
6
     Document d = p.parseDocument();
7
     auto fv = new QueryValidator(d);
8
     auto sv = new SchemaValidator!Schema(d, graphqld.schema);
9
     fv.accept(d);
10
     sv.accept(d);
11
12
     Context con = buildContext(req);
13
     Json ret = graphqld.execute(d, extractVariables(req), con);
14
     res.writeJsonBody(ret);
15
16
```

```
graphqld.setResolver("Employee", "info",
1
    delegate(string name, Json parent, Json args,
2
         ref Context context)
3
    ſ
4
       const id = parent["infoId"].get!long();
5
       EmployeeInfo ei = getEmployeeInfo(id);
6
       Json ret = Json.emptyObject();
7
       ret["data"] = toGraphglJson(ei);
8
       return ret:
9
    });
10
```

- Mostly feature complete
- Some validations are missing
- pprox 17000 lines
- \approx 9000 lines are generated by darser
- ready for use now

Homework

Write a GraphQL backend that uses an excel spreadsheet as a database.

Conclusion

Conclusion

- Spreadsheets are a terrible programming language.
- C++ and Rust are not our main competition on our path to world domination.

Conclusion

- Spreadsheets are a terrible programming language.
- C++ and Rust are not our main competition on our path to world domination.

- We need to learn to use what is there.
- Use D to work smart not hard.
- Do not write the code, write the code that writes the code.
- Look at JS for inspiration.
- GraphQL

The End

- Infographic: C/C++ facts we learned before going ahead with CLion. https://blog.jetbrains.com/clion/2015/07/infographics-cpp-factsbefore-clion/. (Accessed on 04/08/2019).
- [2] Developer Economics. Developer Economics: State of the Developer Nation 15th Edition. https://www.developereconomics.com/reports/state-of-thedeveloper-nation-15th-edition. (Accessed on 04/08/2019).
- [3] Microsoft. Build 2016 Keynote (Day 2). https://www.youtube.com/watch?v=bf0Rr81is6U. (Accessed on 04/08/2019).
- [4] Irish Tech News. Seven reasons why Excel is still used by half a billion people worldwide. https://irishtechnews.ie/seven-reasons-why-excel-isstill-used-by-half-a-billion-people-worldwide/. (Accessed on 04/08/2019).

- [5] Felienne Hermans. GOTO 2016: Pure Functional Programming in Excel by Felienne Hermans. https://www.youtube.com/watch?v=0yKf8TrLUOw. (Accessed on 04/08/2019).
- [6] Marak/faker.js: generate massive amounts of realistic fake data in Node.js and the browser. https://github.com/marak/Faker.js/. (Accessed on 05/03/2019).
- [7] kaleidicassociates/faked: D library to create real fake data. https://github.com/kaleidicassociates/faked. (Accessed on 05/03/2019).
- [8] GraphQL / A query language for your API. https://graphql.org/. (Accessed on 05/03/2019).

- [9] burner/graphqld: A vibe.d library to handle the GraphQL Protocol written in the D Programming Language. https://github.com/burner/graphqld. (Accessed on 05/03/2019).
- [10] burner/Darser: LL1 Parser Generator for D. https://github.com/burner/Darser. (Accessed on 05/03/2019).

Encore

Reappearing UDA Pattern

```
1 struct Employee {
2     @GQLD(
3     Description("The social security number of an employee"),
4     Deprecated(IsDeprecated.yes, "To complex")
5     )
6     SocialSecurityNumber number;
7 }
```

Reappearing UDA Pattern

```
struct Employee {
1
     @GQLD(
2
        Description("The social security number of an employee"),
3
        Deprecated(IsDeprecated.yes, "To complex")
4
5
      )
      SocialSecurityNumber number;
6
7 }
   enum IsDeprecated {
9
     undefined,
10
11
     no.
12
     yes
13
   }
14
   struct GQLDData {
15
     Description desc;
16
     Deprecated depre;
17
  }
18
```

Reappearing UDA Pattern

```
struct GQLDData {
1
     Description desc;
2
     Deprecated depre;
3
4 }
5
   GQLDData GQLD(Args...)(Args args) {
6
     GQLDData ret;
7
     static foreach(mem; __traits(allMembers, GQLDData)) {
8
       static foreach(arg; args) {
9
         static if(is(typeof(__traits(getMember, ret, mem)) ==
10
              typeof(arg)))
11
         ł
12
            traits(getMember, ret, mem) = arg;
13
         }
14
       }
15
     }
16
17
     return ret:
```

```
Json ret = Json.emptyObject();
1
2
   string typename = ...;
   l: switch(typename) {
3
     static foreach(type; collectTypes!(T)) {{
4
       case typeToTypeName!(type): {
5
         ret["data"] = typeToJson!(type)();
6
         break 1;
7
      }
8
    }}
9
     default: break;
10
  7
11
  return ret;
12
```

Collecting all Referenced Types

```
alias allTypes = collectTypes!Schema;
1
2
   template collectTypesImpl(Type) {
3
     import graphql.uda;
4
     static if(is(Type : GQLDCustomLeaf!F, F)) {
5
       alias collectTypes Impl= AliasSeq!(Type);
6
     } else static if(is(Type == interface)) {
7
       alias RetTypes = AliasSeq!(collectReturnType!(Type,
8
           __traits(allMembers, Type)));
9
       alias ArgTypes = AliasSeq!(collectParameterTypes!(Type,
10
           __traits(allMembers, Type)));
11
       alias collectTypesImpl = AliasSeq!(Type, RetTypes,
12
           ArgTypes, InterfacesTuple!Type);
13
14
     } else static if(is(Type == union)) {
15
16
       alias collectTypesImpl = AliasSeq!(Type, InheritedClasses!Type);
     } else static if(is(Type : Nullable!F, F)) {
17
       alias collectTypesImpl = .collectTypesImpl!(F);
18
19
     . . .
```

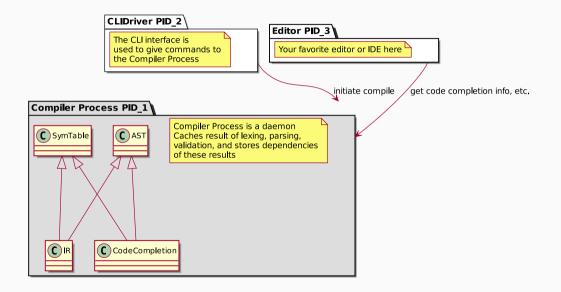
Type trinary expression

```
template InheritedClass(T) {
1
     import std.meta : staticMap, AliasSeq, NoDuplicates;
2
     import std.traits : Select;
3
4
5
     alias getInheritedFields() = staticMap!(.InheritedClass, FieldTypeTuple!T);
     alias ftt = Select!(is(T == union), getInheritedFields, AliasSeq);
6
7
     alias getBaseTuple() = staticMap!(.InheritedClass, BaseClassesTuple!T);
8
     alias clss = Select!(is(T == class), getBaseTuple, AliasSeq);
9
10
     alias getInter() = staticMap!(.InheritedClass, InterfacesTuple!T);
11
     alias inter = Select!(is(T == class) || is(T == interface).
12
         getInter,
13
         AliasSeq
14
15
       );
16
17
     alias InheritedClass = NoDuplicates!(AliasSeq!(ftt!(), clss!(), inter!()));
  3
18
```

- \approx 7000 lines
- \approx 8 seconds build time

• Traditional compiler pipeline design is dated

- Traditional compiler pipeline design is dated
- We have practically unlimited memory
- Recreating the AST, IR, and ASM on every compile is extremely wasteful
- Why does code-completion and the compiler use different frontends





- Darser is a recursive descent parser generator for LL(1) grammars
- It also generates the AST and a default Visitor
- Not at CT, but as a pre-build step



- Darser is a recursive descent parser generator for LL(1) grammars
- It also generates the AST and a default Visitor
- Not at CT, but as a pre-build step

- It generates "good" error messages
- Not just a generic Node types, but names that reflect the grammar
- Inheriting from the default Visitor is trivial and powerful
- Used right now by GraphQLD

I am out of Slides