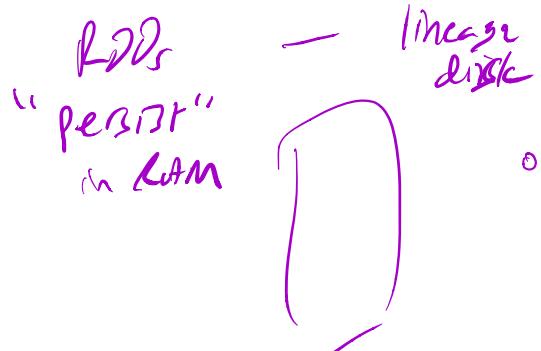


SparkSQL

Spark - reduces I/O

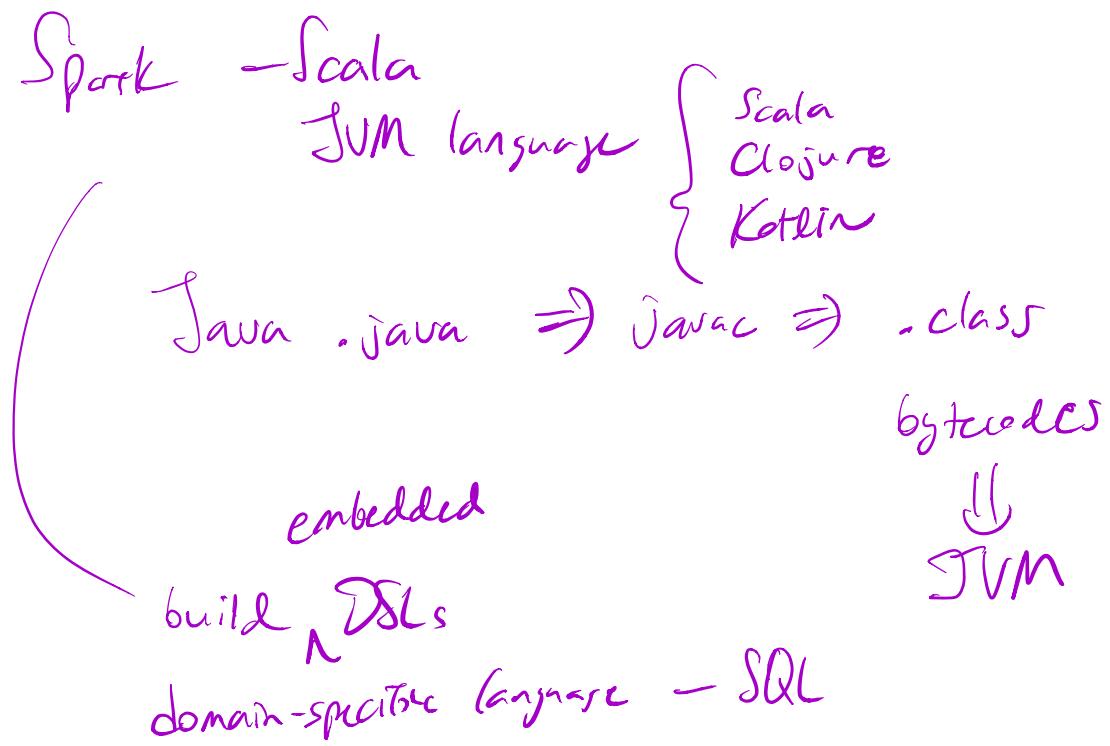
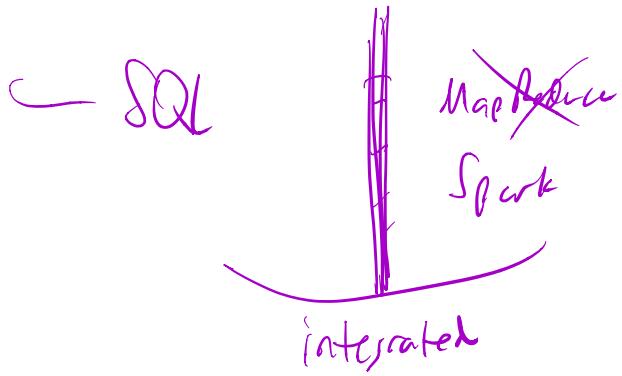


reduce I/O
- faster

~ UDFs, black boxes
exposing "the guts" of UDFs
so that we can optimize
SQL + UDFs

eval(s) = ?

Java mostly static
- "reflection"
- class loaders } dynamiz



quasiquote

square

#define $\text{do}(x)$ $((x)*(x))$

:

square

$\text{do}(12) \rightarrow ((12)*(12))$

int square(int x){
 return x*x;
}

} eliminates

function call overhead

C++ "inline"

define while if
preprocessor

square(x)

source → macros → compiler
expanded

#define square(x) x*

C/C++
~~assert(x > 0);~~

-DNDEBUG

assert(isEqual(x, sort(x)));

Java

|||
NO MACROS
... AND

safe
useful

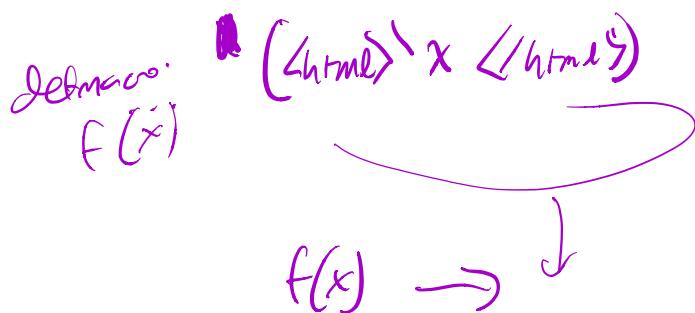
C
ANYTHING
GOES

AUWWWW

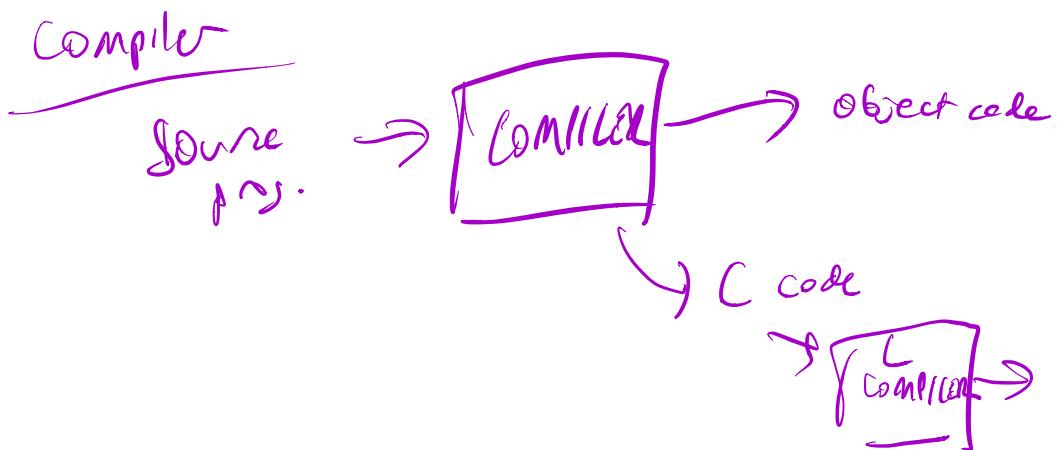
"hygienic macros"

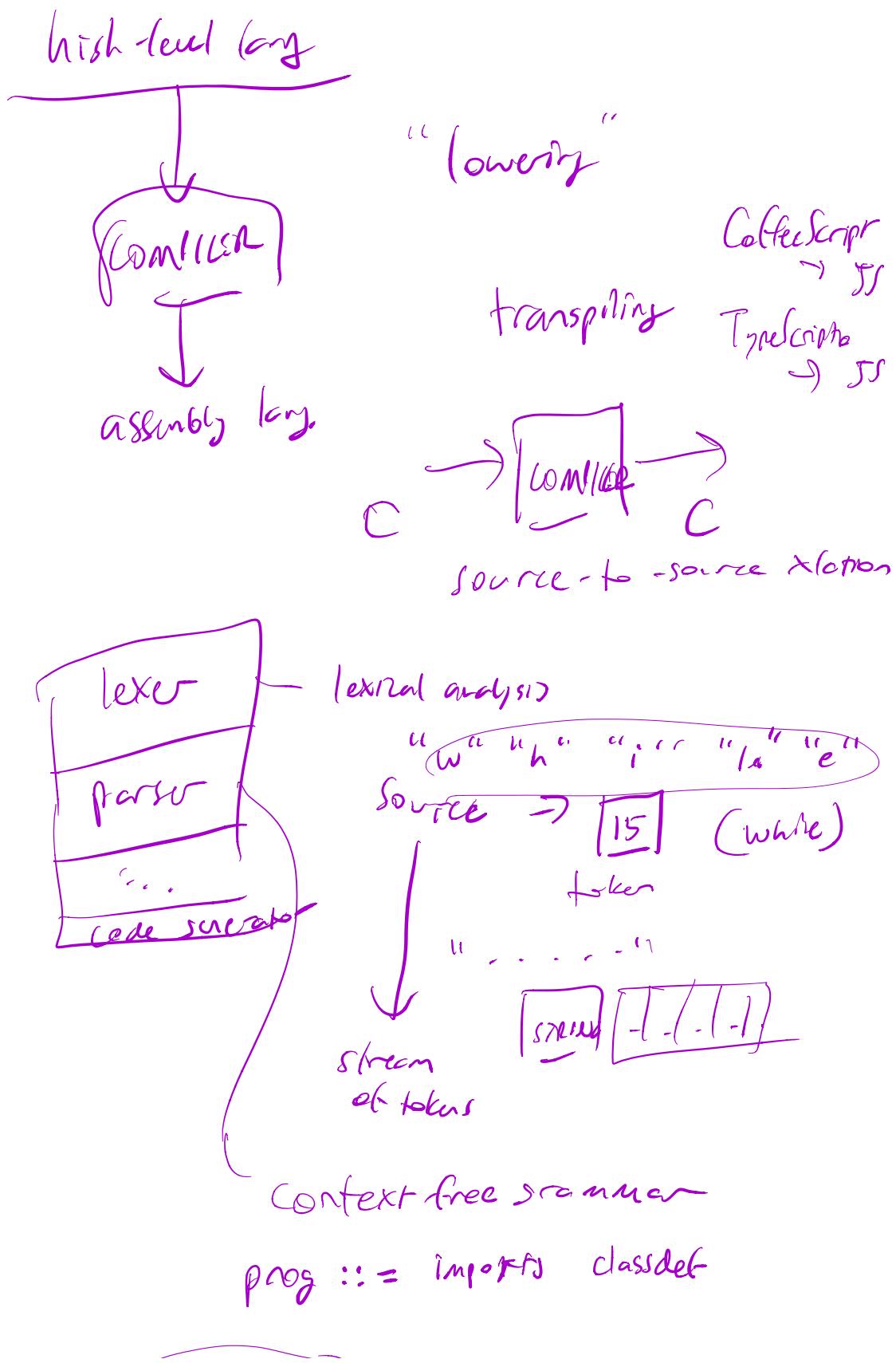
`#define foo(x, x*y)`

'() quoting

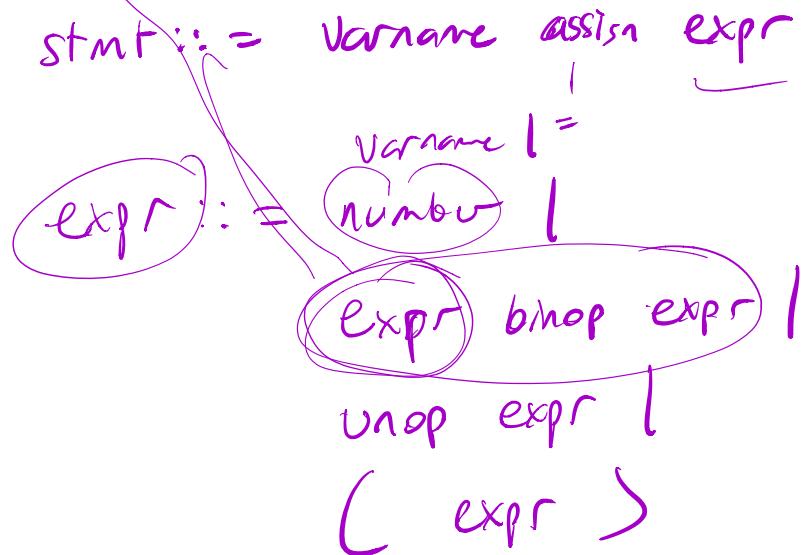


~~while code it~~





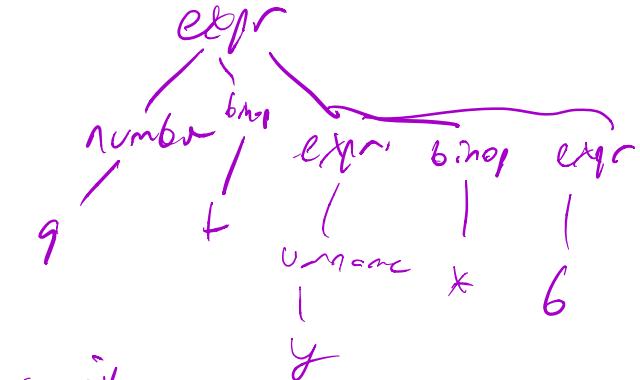
$$x = 9 + (y * 6)$$



stmt ::= varname assign expr

If

Abstract
syntax tree



compiler

"intermediate
representation"

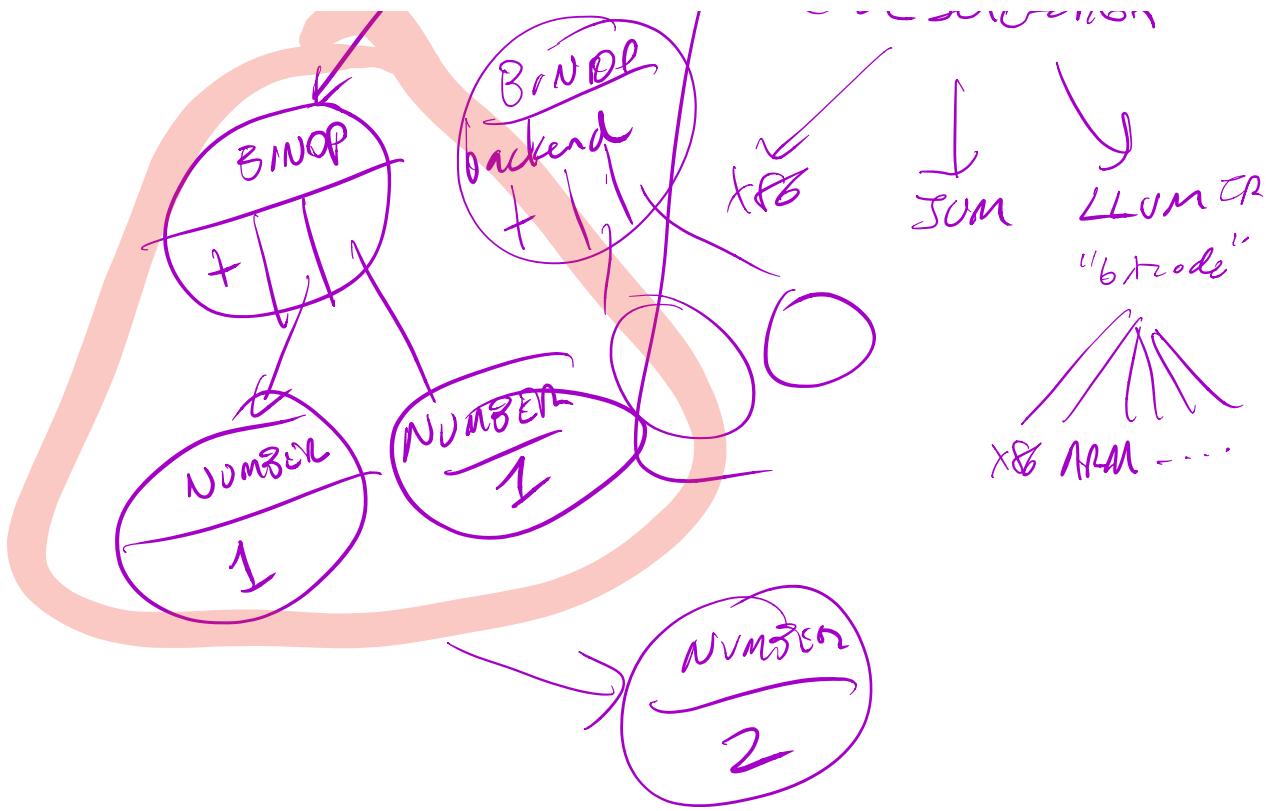
three-value code
SSA

$$x = 1 + 1 \oplus 1 + 1$$

BINOP

+

Code gen. - al...

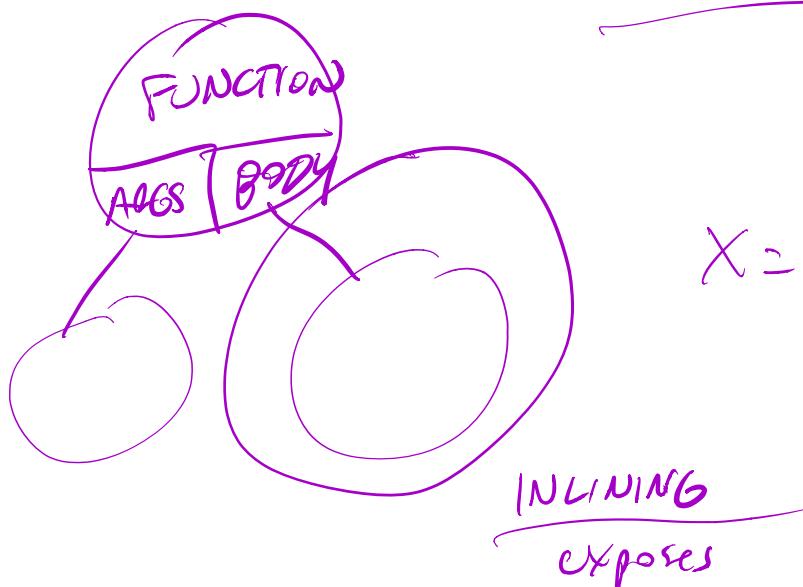


$X = 4$

CONSTANT or CONSTANT

\Rightarrow CONSTANT

constant folding



$$X = f(\underbrace{g(x)}, \underbrace{h(x)})$$

$$g(x) = \\ \text{return } x + 2;$$

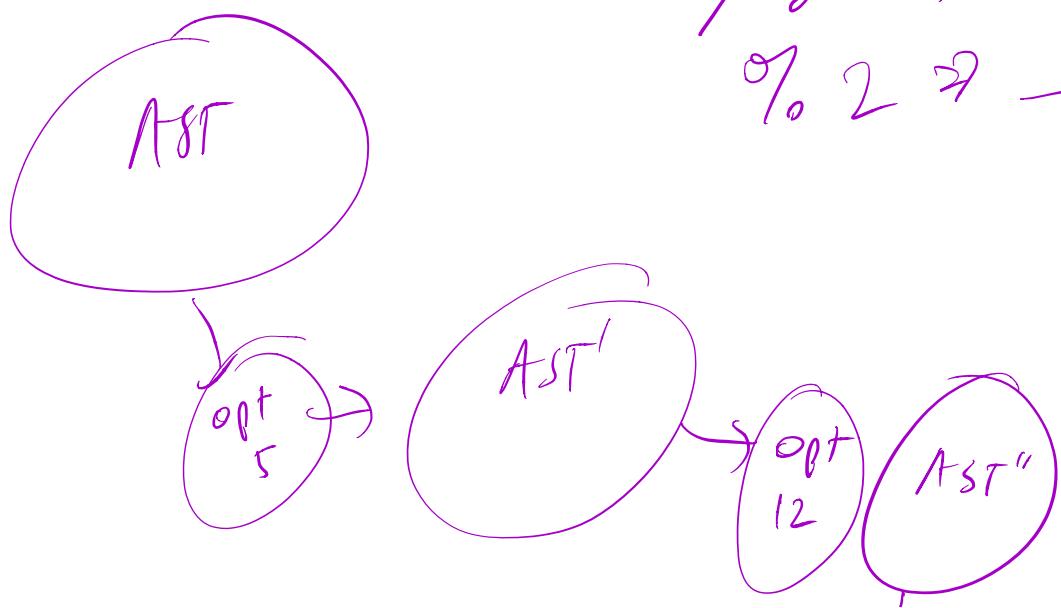
$X = 12;$
 $Y = 100;$
 $Z = \text{download all websites \& concatenate them}$
 $\text{Sps... write}(\text{'HelloWorld'});$

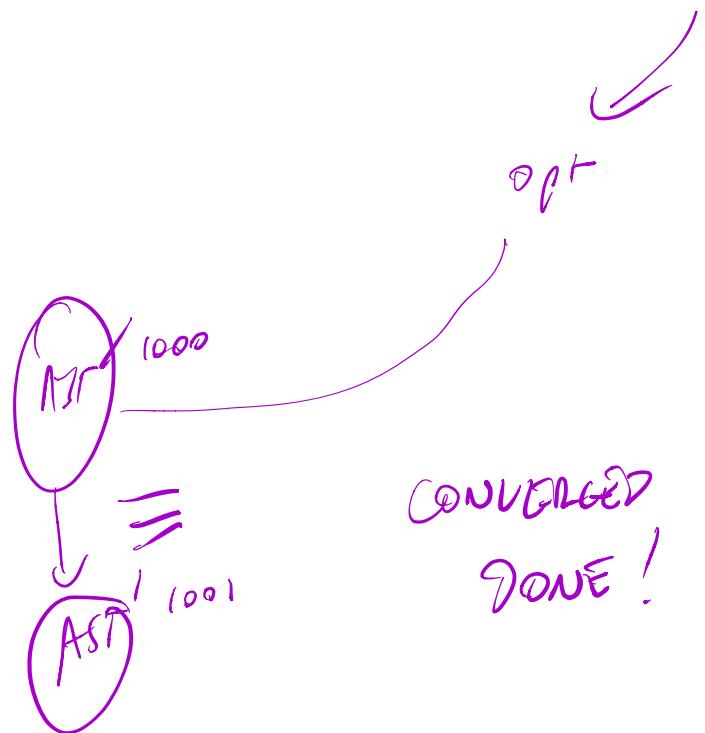
Optimization opportunities

Dead code elimination
 liveness analysis
 class Foo {
 x
 y
 z
 x

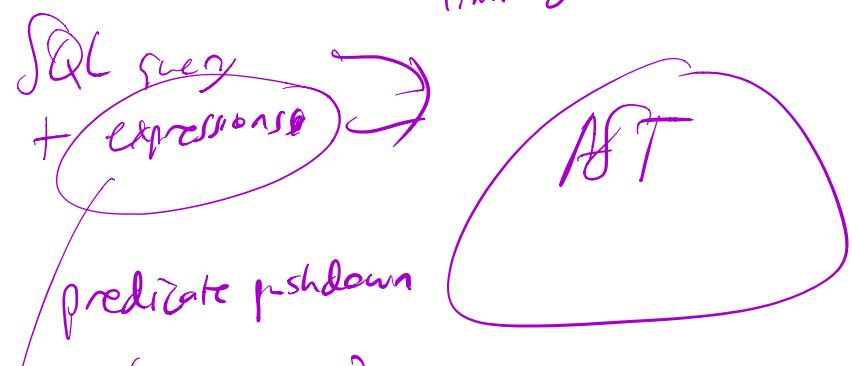
strength reduction

$/2 \Rightarrow \gg$
 $\%2 \Rightarrow -$



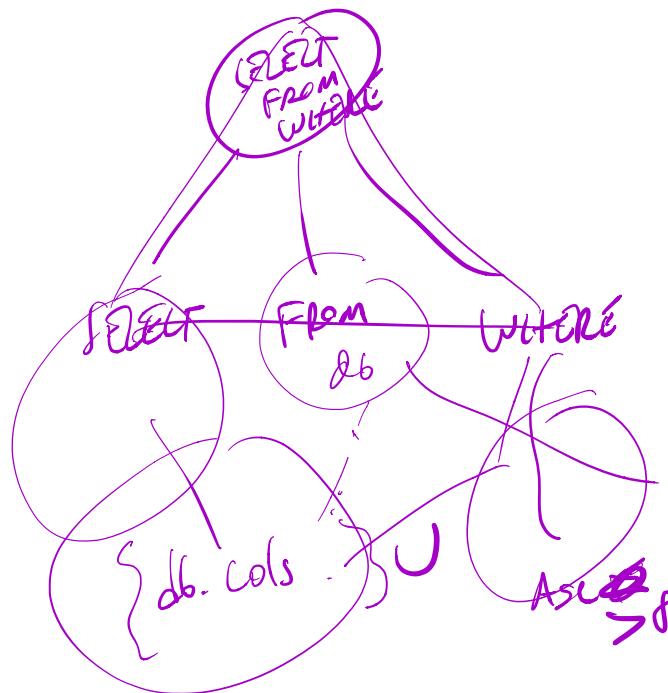
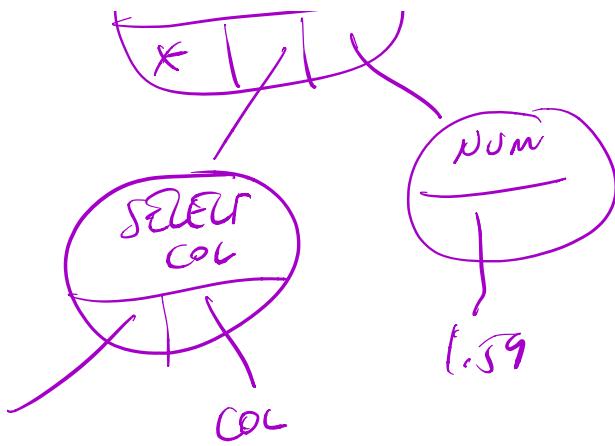


SparkSQL - optimize SQL queries
optimize UDFs
 └── limited expression language



⊗ $x[\text{"Salon"}] \times 1.59$

BINOP

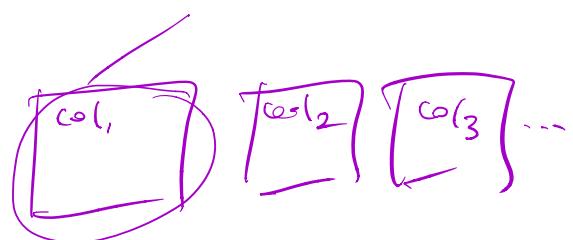


SparkSQL

"data frames"

Columnar

Catalyst



```
def compile(node: Node): AST = node match {
  case Literal(value) => q"$value"
  case Attribute(name) => q"row.$name"
  case Add(left, right) => q"$left + $right"
}
```

• • • • L PLACEMENT / T
\$compile(nest)"

self-hosted