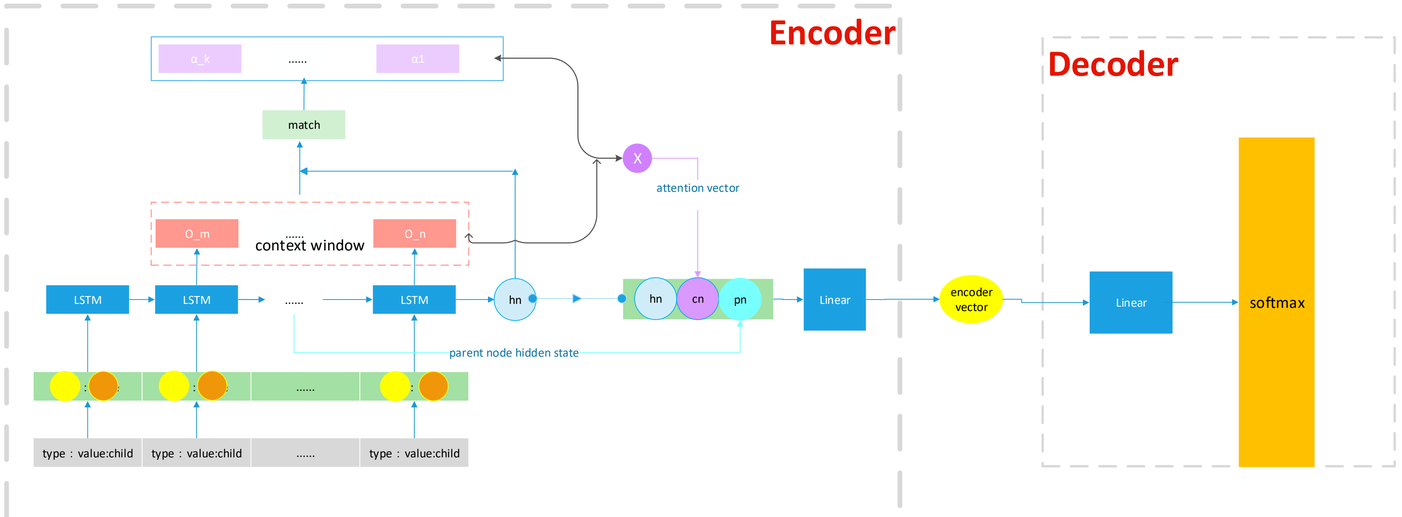
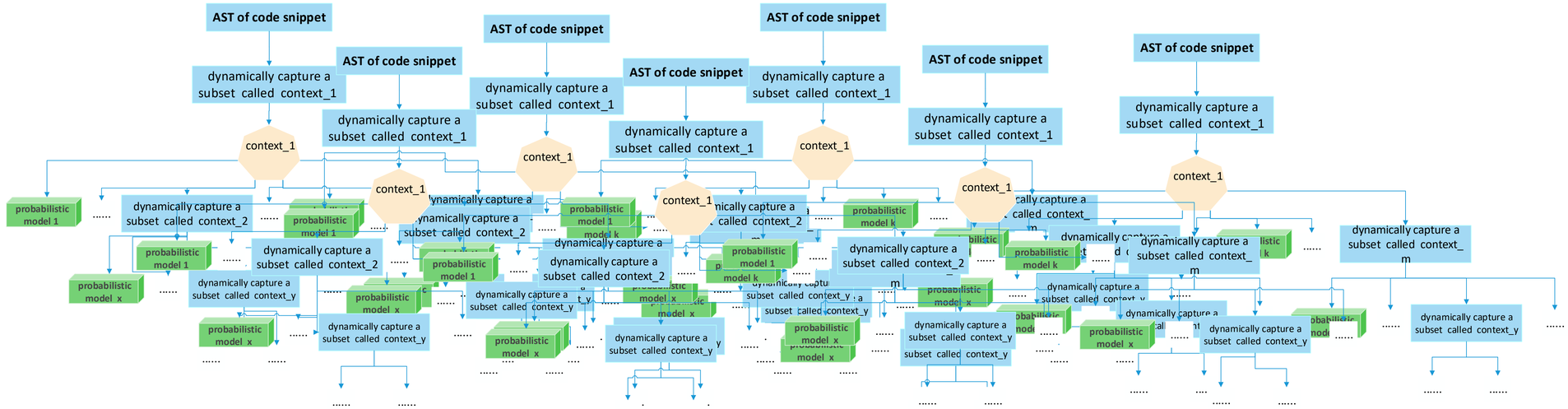
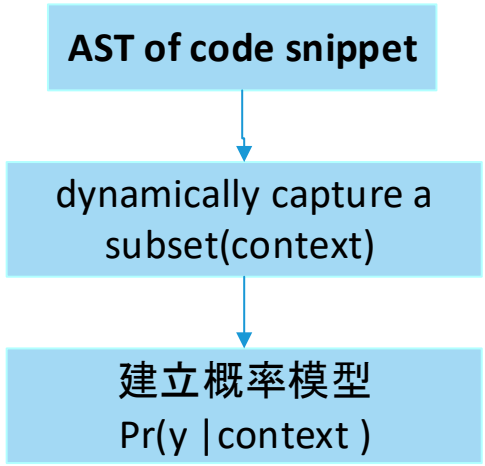
type : value : children	type : value : children	.....	type : value

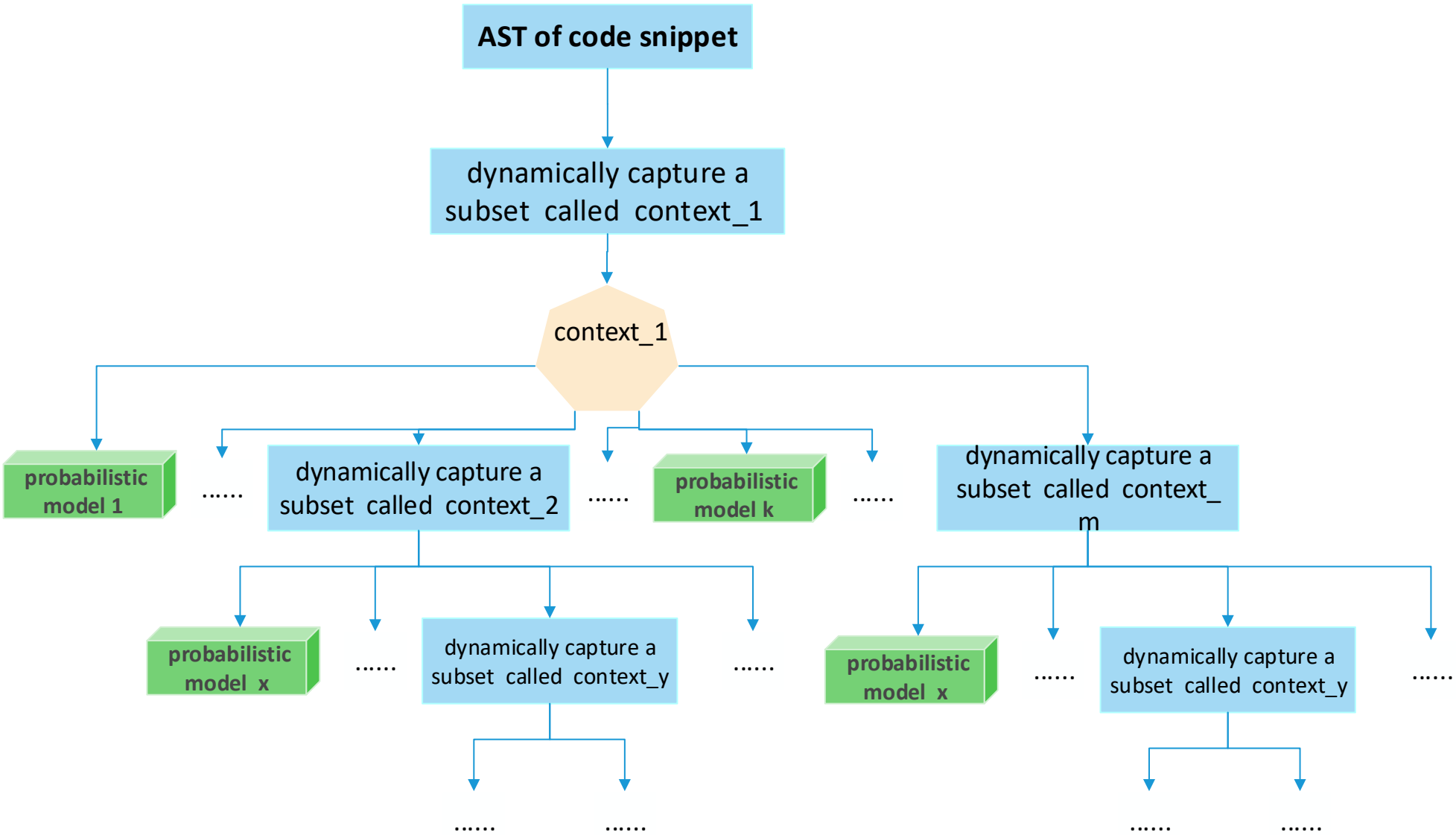
x = 7

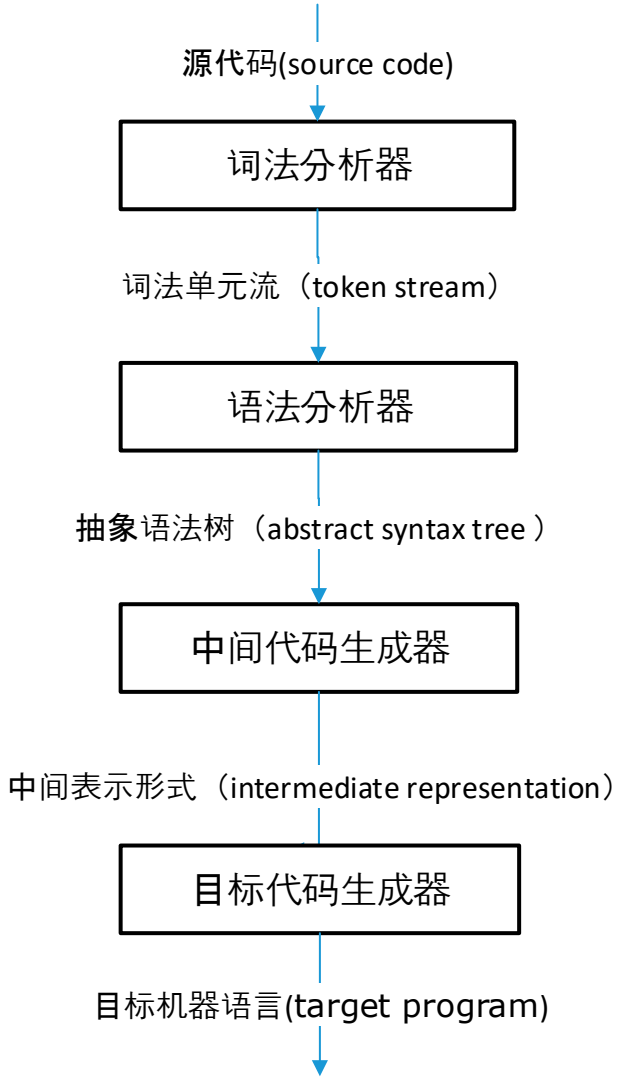
{"type":"Module","children":[1]}	{"type":"Assign","children":[2,3]}	{"type":"NameStore","value":"x"}	{"type":"Num","value":"7"}

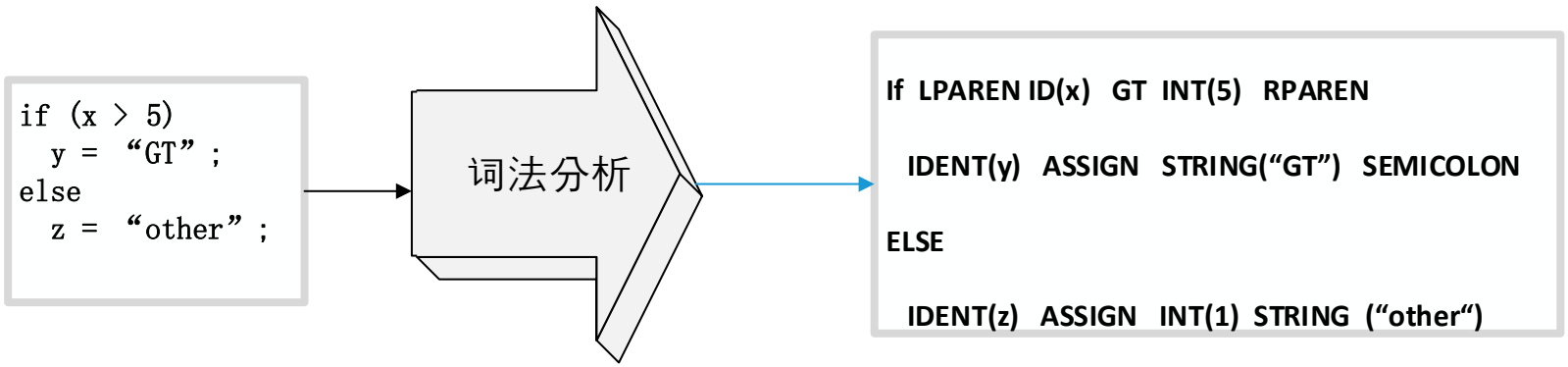

Encoder
Decoder
α_k
.....
α1
match
X
attention vector
O_m
.....
context window
O_n
LSTM
LSTM
.....
LSTM
hn
hn
cn
pn
Linear
encoder vector
Linear
softmax
parent node hidden state
.....
type : value:child
type : value:child
.....
type : value:child

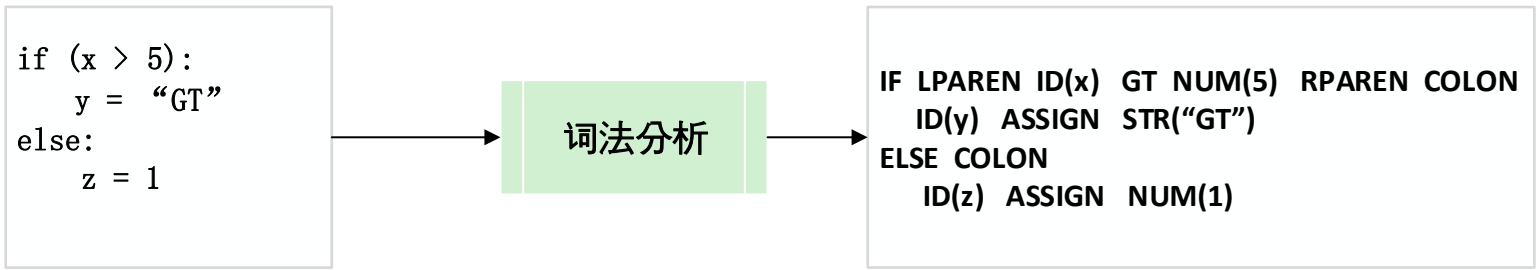

AST of code snippet
dynamically capture a subset called context_1
context_1
dynamically capture a subset called context_2
probabilistic model 1
probabilistic model k
dynamically capture a subset called context_y
probabilistic model x
dynamically capture a subset called context_ m

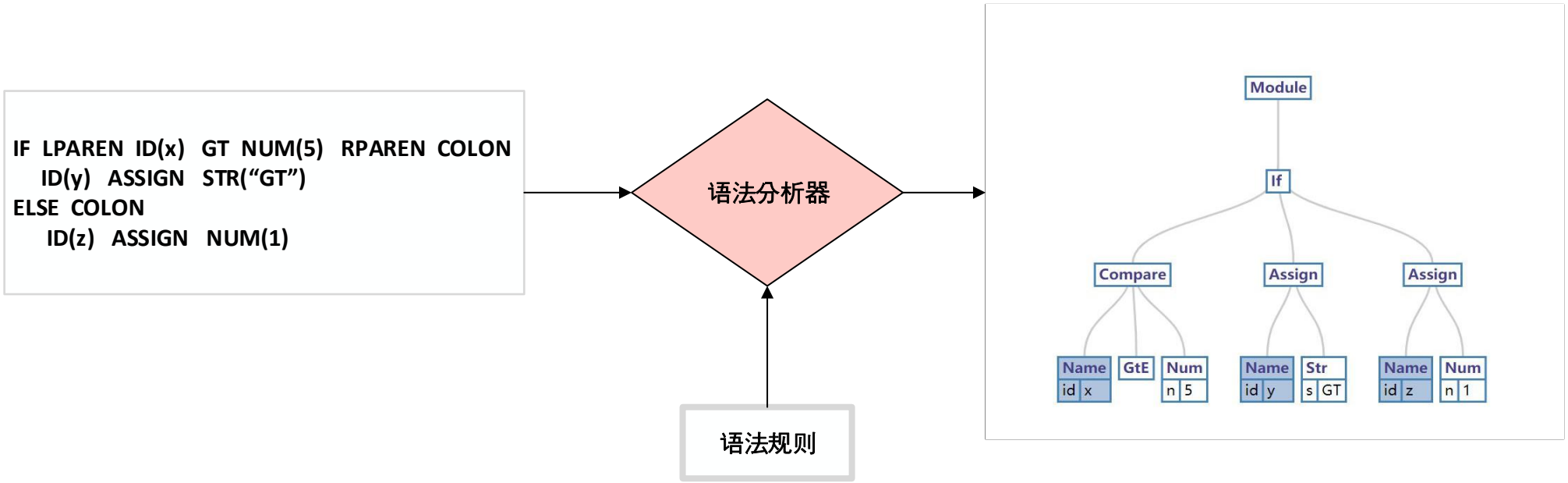

AST of code snippet
dynamically capture a subset(context)
建立概率模型
Pr(y |context)


AST of code snippet
dynamically capture a subset called context_1
context_1
probabilistic model 1
.....
dynamically capture a subset called context_2
.....
probabilistic model k
.....
dynamically capture a subset called context_m
probabilistic model x
.....
dynamically capture a subset called context_y
.....
.....
.....
probabilistic model x
.....
dynamically capture a subset called context_y
.....
.....
.....

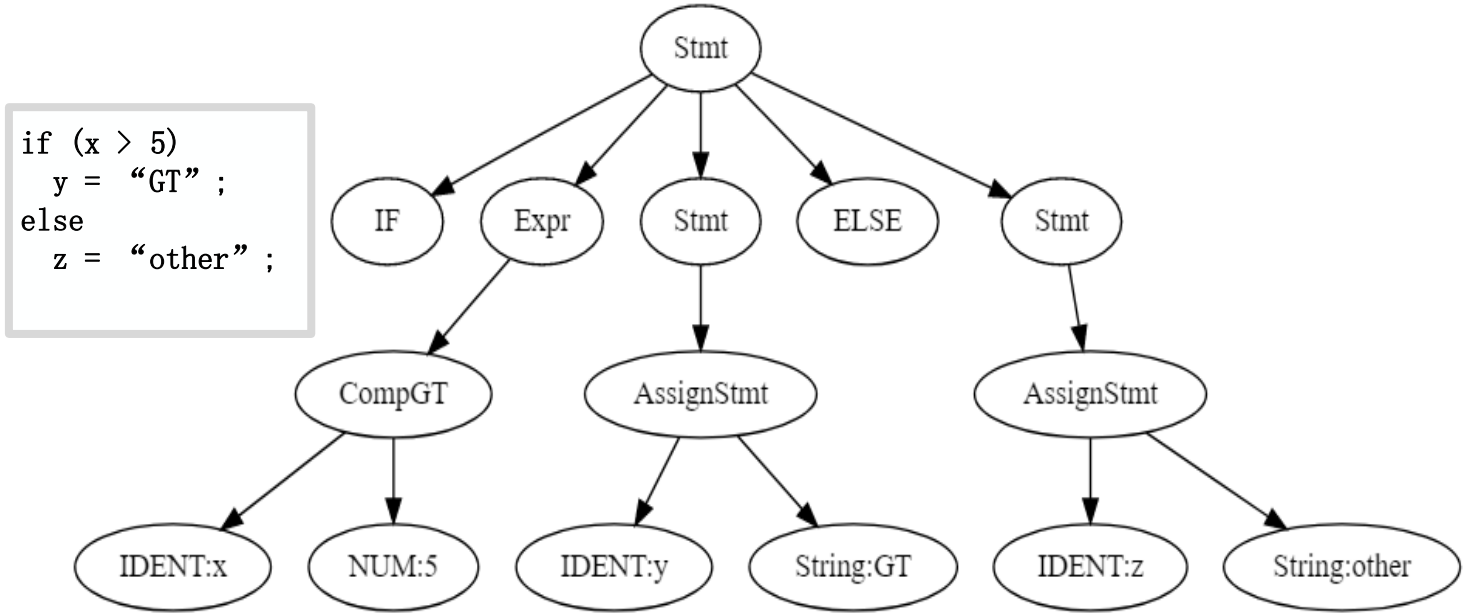

源代码(source code)
词法分析器
词法单元流（token stream）
语法分析器
抽象语法树（abstract syntax tree ）
中间代码生成器
中间表示形式（intermediate representation）
目标代码生成器
目标机器语言(target program)

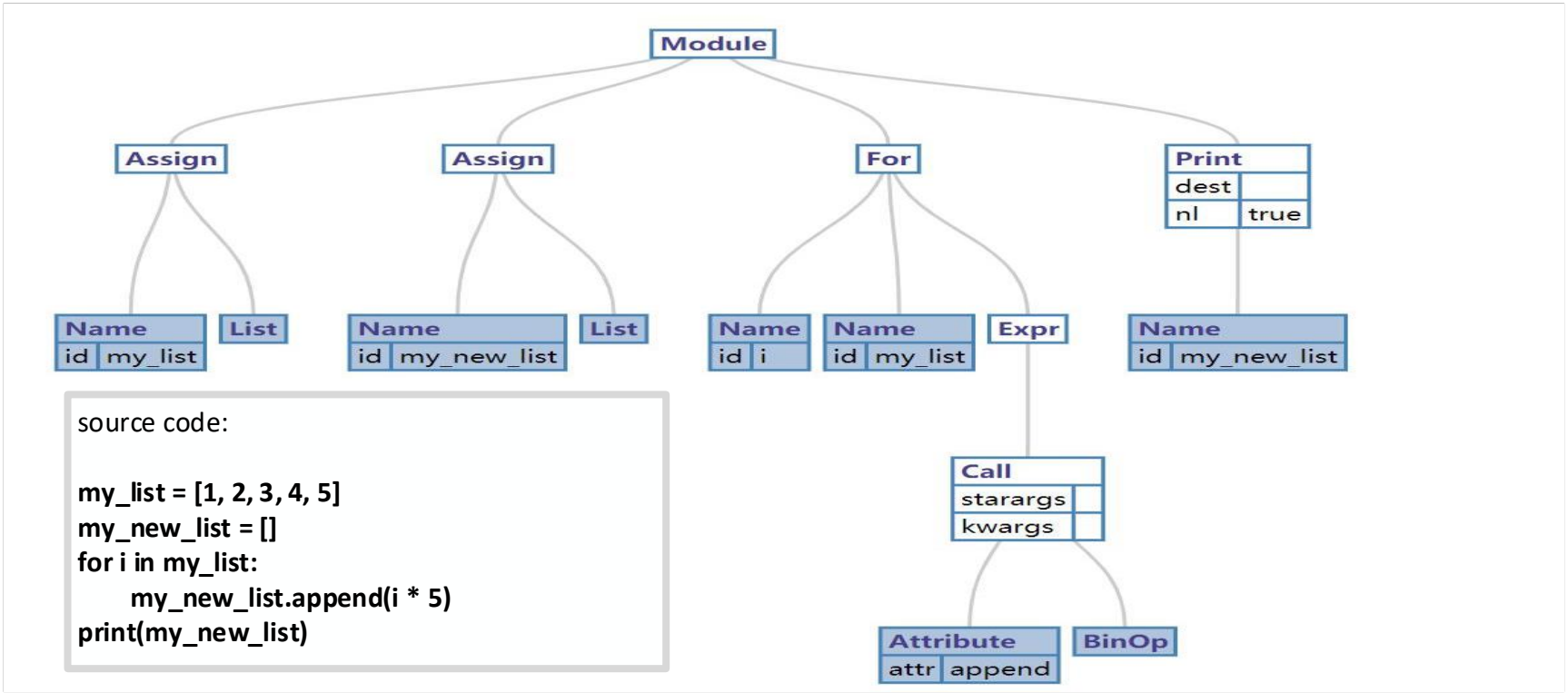

if (x > 5)
y = “GT” ;
else
z = “other” ;
词法分析
If LPAREN ID(x) GT INT(5) RPAREN
IDENT(y) ASSIGN STRING(“GT”) SEMICOLON
ELSE
IDENT(z) ASSIGN INT(1) STRING (“other“)


if (x > 5):
y = “GT”
else:
z = 1
词法分析
IF LPAREN ID(x) GT NUM(5) RPAREN COLON
ID(y) ASSIGN STR(“GT”)
ELSE COLON
ID(z) ASSIGN NUM(1)


IF LPAREN ID(x) GT NUM(5) RPAREN COLON
ID(y) ASSIGN STR("GT")
ELSE COLON
ID(z) ASSIGN NUM(1)
语法分析器
语法规则
Module
If
Compare
Assign
Assign
Name
id x
GtE
Num
n 5
Name
id y
Str
s GT
Name
id z
Num
n 1

```
if (x > 5)
  y = “GT” ;
else
  z = “other” ;
```





source code:

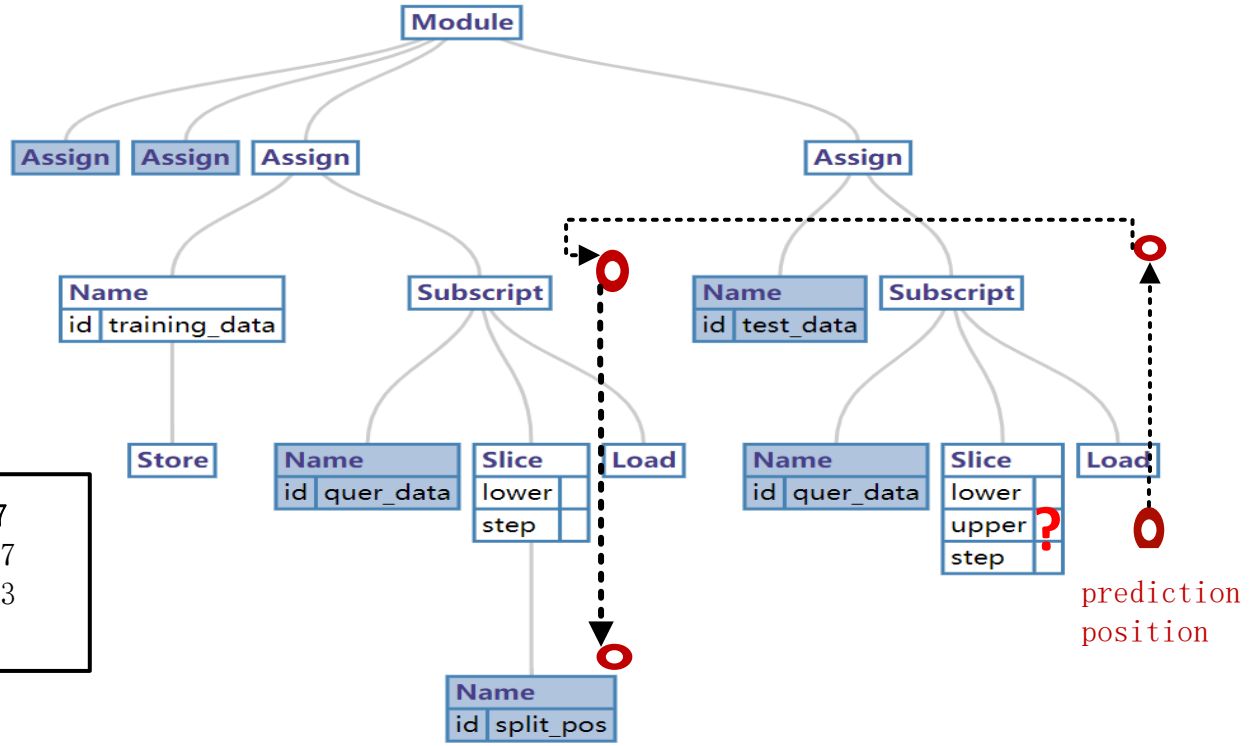
```
my_list = [1, 2, 3, 4, 5]
my_new_list = []
for i in my_list:
    my_new_list.append(i * 5)
print(my_new_list)
```

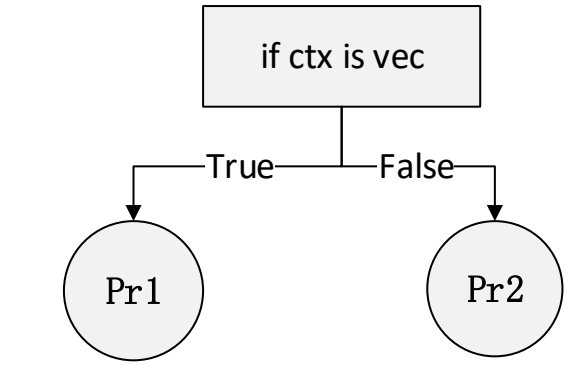
```
...
split_pos = int(len(quer_data) * 0.7)

training_data = quer_data[:split_pos]

test_data = quer_data[?:]
...
```

split_pos	0.87
i	0.07
len	0.03
...	





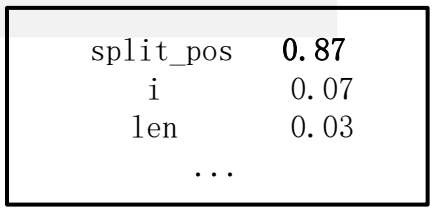
Conditional Pr2
numpy ->np
tensorflow ->tf

Unconditional Pr2
np 50%
tf 50%

```
...
split_pos = int(len(quer_data) * 0.7)

training_data = quer_data[:split_pos]

test_data = quer_data[?:]
...
```



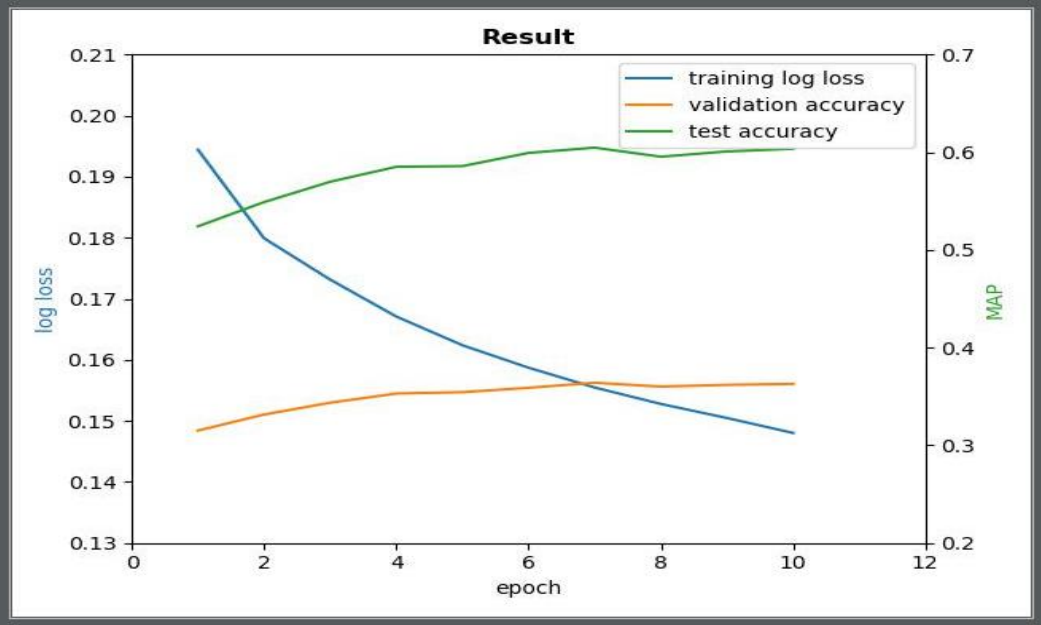
```
start= time.time()
. . .
MAP = eval(tree, data, vec)

now = time.time()

min, sec = divmod((now - start), 60)

hour, min = divmod( , )
...
```

?



5774221079]
18104267]
444956]
2106]

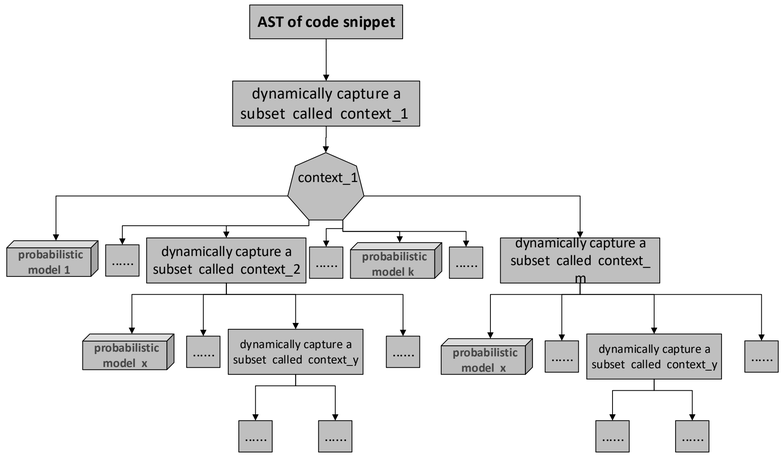
[Epoch:10] train_loss:0.148025 val_loss:0.170486 | time spend: 423.864023
train loss [0.19442062096255167, 0.17990774575982774, 0.1731374791775431, 0.16709070628285408, 0.16238965774221079, 0.15870868202745914, 0.15546583086252214, 0.15277694323637656, 0.15047304502044406, 0.14802473999496016]
eval loss [0.1849578297416369, 0.1779695038954417, 0.17367957796156405, 0.17148049709796906, 0.17058431318104267, 0.1687246495783329, 0.16729239653348924, 0.16913055136700472, 0.1699556341846784, 0.1704856427371502]
eval MAP [0.3151083333333429, 0.3315380555555625, 0.3436008333333392, 0.35304083333333636, 0.35440194444444956, 0.35887833333333713, 0.3645944444444804, 0.36011583333333697, 0.3618941666666712, 0.3628622222222258]
test MAP [0.52416611111110955, 0.5489549999999841, 0.5698305555555415, 0.5851499999999893, 0.5858197222222106, 0.5993441666666545, 0.604734444444366, 0.5954488888888786, 0.600843055555548, 0.6035763888888787]
Epoch 11/100

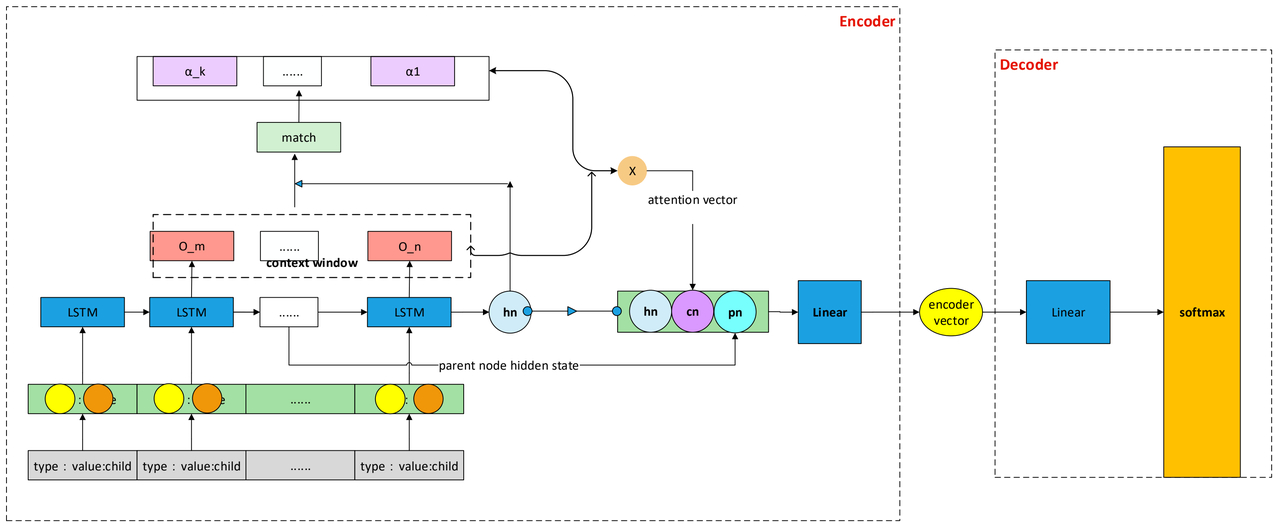
[Epoch:11] train_loss:0.145954 val_loss:0.171683 | time spend: 465.963525
Epoch 12/100

[Epoch:12] train_loss:0.144762 val_loss:0.172393 | time spend: 507.993734
Epoch 13/100

[Epoch:13] train_loss:0.143241 val_loss:0.172773 | time spend: 550.139675
Epoch 14/100

[Epoch:14] train_loss:0.141747 val_loss:0.173340 | time spend: 592.274313
Epoch 15/100


AST of code snippet
dynamically capture a subset called context_1
context_1
probabilistic model 1
.....
dynamically capture a subset called context_2
.....
probabilistic model k
.....
dynamically capture a subset called context_m
probabilistic model x
.....
dynamically capture a subset called context_y
.....
.....
.....
probabilistic model x
.....
dynamically capture a subset called context_y
.....
.....
.....


Encoder
α_k
.....
α1
match
X
attention vector
O_m
.....
O_n
context window
LSTM
LSTM
.....
LSTM
hn
hn
cn
pn
Linear
parent node hidden state
type : value:child
type : value:child
.....
type : value:child
encoder vector
Decoder
Linear
softmax