

U

O

W

Software Requirements, Specifications and Formal Methods

A/Prof. Lei Niu



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Schemas and schema calculus



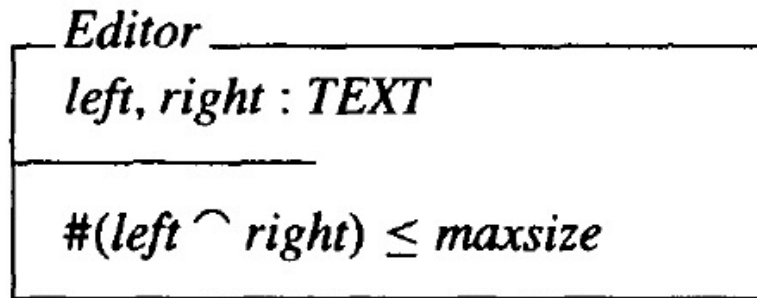
Steps for Modelling with Z Specification

1. Define the data type
 1. free type
 2. basic type
2. Define the axiomatic descriptions
 1. global variables
 2. global constrains
3. Define the system state schema
 1. local variables
 2. system state invariants
4. Initialise the system state schema
 1. Initialise all local variables and global variables
5. Create full operation schema
 1. successful scenarios (when all pre-conditions are true)
 2. non-successful scenarios (when each pre-condition is false)
 3. combine the successful and non-successful scenarios

Inside the schema boxes

We will use our simple editor example to explain the notation to write a single schema box

- Define type: $[CHAR]$
 $TEXT == seq\ CHAR$
- Define state schema: the simple editor's state:



- This schema says that the editor consists of two texts named left and right
- The size of the document never exceeds *maxsize*
- Schema names usually begin with an initial capital letter followed by lowercases letters

Schema inclusion

Then we define the editor's initial state

<i>Init</i>
<i>Editor</i>
$left = right = \langle \rangle$

- The editor starts up with an empty document
- The *Init* schema includes all the declarations and predicates in the state schema *Editor*.
- A longer version can be

<i>Init</i>
$left, right : TEXT$
$\#(left \frown right) \leq maxsize$ $left = right = \langle \rangle$

Operation schema

Define the full operation schema: Z uses operation schemas to model changes of state

- Axiomatic definition can be used anywhere

| *printing* : $\mathbb{P}CHAR$

<i>Insert</i>
$\Delta Editor$
$ch? : CHAR$
$ch? \in printing$
$left' = left \frown \langle ch? \rangle$
$right' = right$

- Δ tells us that Insert is an operation that changes the state of Editor
- ? Tells use that $ch?$ is the input variable
- The prime ' tells use the state after the operation

Operation schema

We can decorate the whole schema with the prime.

- Defines a new schema with the prime
- All variable names in the new schema are marked with the prime

<i>Editor'</i>
<i>left', right' : TEXT</i>
<i>$\#(left' \frown right') \leq maxsize$</i>

- No prime mark on a global variable *maxsize*, because it is not a state variable in schema *Editor*

Operation schema

- The Δ naming convention is just an abbreviation for the schema that include both the unprimed “before” state and the primed “after” state.
- Δ is not an operator

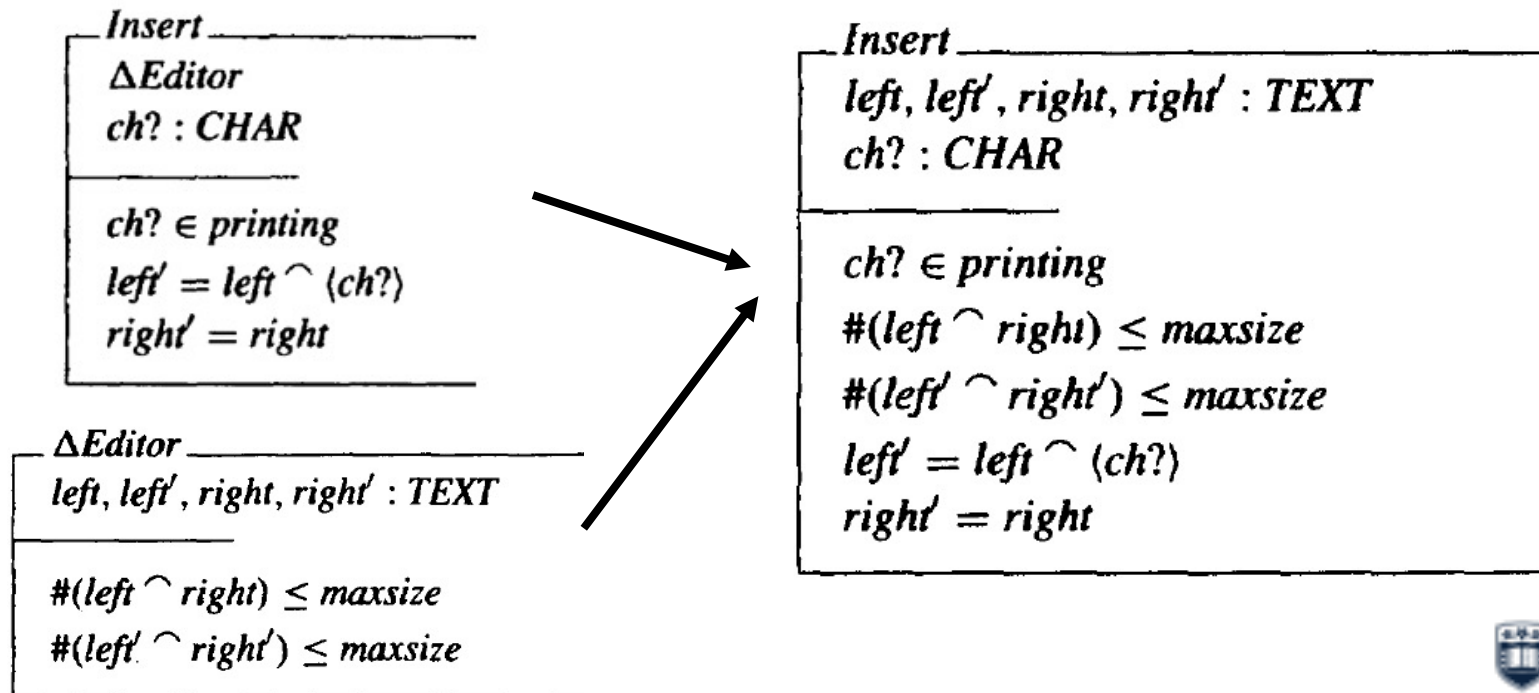
Δ Editor
Editor
Editor'

- When we expand it, we obtain

Δ Editor
$left, left', right, right' : TEXT$
$\#(left \frown right) \leq maxsize$
$\#(left' \frown right') \leq maxsize$

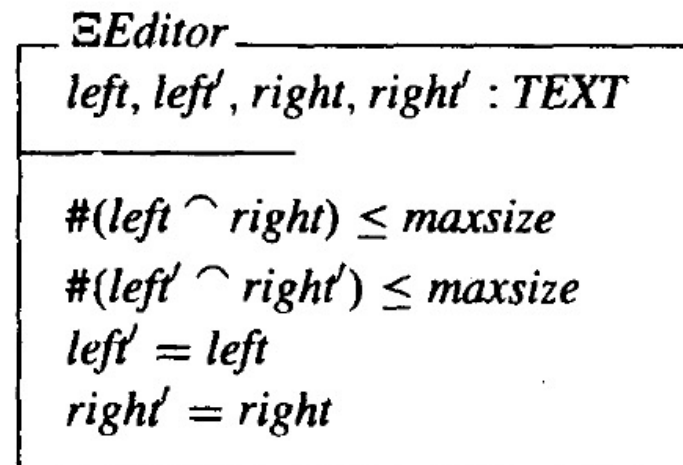
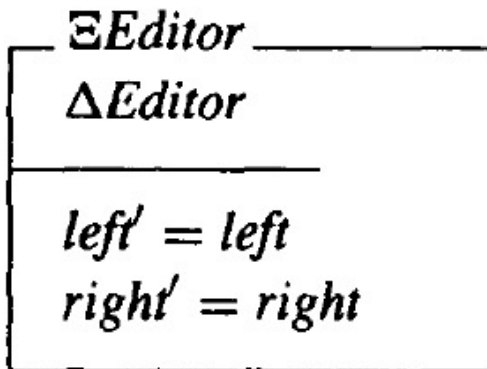
Operation schema

- In schemas, the repeated predicate is always true: it holds before and after any operation
- The predicate is an invariant
- So we can expand the insert operation by replacing the $\Delta Editor$ with its full text



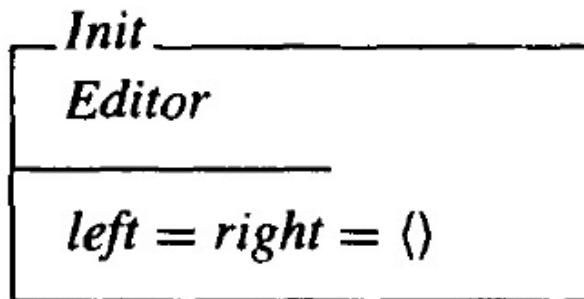
Operation schema

- The Ξ symbol indicates an operation where the state does not change
- The operation only consumes inputs or produce outputs
- Ξ operation is another name convention.



Vertical and horizontal schema format

- Z provides different ways to notate schemas
 - Vertical format



- Horizontal format

Init $\hat{=}$ [*Editor* | *left = right = ⟨⟩*]

In the horizontal format, the schema name appears to the left of the definition system, and the schema body is enclosed within square brackets, with a vertical bar separating the declaration and predicates.

Vertical and horizontal schema format

- Vertical format

<i>Insert</i>
$\Delta Editor$
$ch? : CHAR$
$ch? \in printing$
$left' = left \frown \langle ch? \rangle$
$right' = right$

<i>Insert</i>
$\Delta Editor; ch? : CHAR$
$ch? \in printing \wedge left' = left \frown \langle ch? \rangle \wedge right' = right$

- Horizontal format

$$Insert \hat{=} [\Delta Editor; ch? : CHAR \mid ch? \in printing \wedge left' = left \frown \langle ch? \rangle \wedge \dots]$$

Schema conjunction and disjunction

Schema conjunction combines the predicates using the logical connective and (\wedge)

For example, 12 divided by 5 yields quotient 2 and remainder 2.

<i>Quotient</i>
$n, d, q, r : \mathbb{N}$
$d \neq 0$
$n = q * d + r$

Any problem?

Schema conjunction and disjunction

Schema conjunction combines the predicates using the logical connective and (\wedge)

For example, 12 divided by 5 yields quotient 2 and remainder 2.

Quotient
$n, d, q, r : \mathbb{N}$
$d \neq 0$ $n = q * d + r$

- $12 = 2 * 5 + 2$ (seems ok, but not good enough) because
- $12 = 0 * 5 + 12$ (i.e., 12 divided by 5 gives quotient 0 and remainder 12) ✗

Schema conjunction and disjunction

So we shall also say that the remainder is less than the divisor

<i>Remainder</i>
$r, d : \mathbb{N}$
$r < d$

Finally, we can form the complete specification using the schema conjunction operator

$$\textit{Division} \hat{=} \textit{Quotient} \wedge \textit{Remainder}$$

Schema conjunction and disjunction

A typical Z style:

- Define requirements separately
- Use the schema conjunction operator to combine the requirements



Schema conjunction and disjunction

- Schema disjunction combines the predicates using the logical connective or.
- We use disjunction to handle separate cases, especially errors and other exceptional conditions
- For example, our Division schema is partial: It doesn't say what happens when the divisor d is zero.
- Let's define a DivideByZero schema first.

Schema conjunction and disjunction

<i>DivideByZero</i>
$d, q, r : \mathbb{N}$
$d = 0 \wedge q = 0 \wedge r = 0$

- Then we join the normal and exceptional cases to describe the total operation

$$T_Division \hat{=} Division \vee DivideByZero$$

- The specification can be expressed as a single schema box

<i>T_Division</i>
$n, d, q, r : \mathbb{N}$
$(d \neq 0 \wedge r < d \wedge n = q * d + r) \vee$ $(d = 0 \wedge r = 0 \wedge q = 0)$

Schema conjunction and disjunction

- Of course, we can combine conjunction and disjunction operators

$$T_Forward \hat{=} Forward \vee (EOF \wedge RightArrow \wedge \exists Editor)$$

Forward

left, right, left', right' : TEXT
ch? : CHAR

ch? = right_arrow
right ≠ {}
 $\#(left \frown right) \leq maxsize$
 $\#(left' \frown right') \leq maxsize$
left' = left \frown (head right)
right' = tail right

EOF

left, right : TEXT

$\#(left \frown right) \leq maxsize \wedge right = \{\}$

RightArrow

ch? : CHAR

ch? = right_arrow

∃Editor

left, right, left', right' : TEXT

$\#(left \frown right) \leq maxsize$
 $\#(left' \frown right') \leq maxsize$
left' = left \wedge right' = right



Schema conjunction and disjunction

Merging the declarations, combining the predicates, and simplifying, we obtain a full specification of $T_Forward$.

$T_Forward$

$left, right, left', right' : TEXT$

$ch? : CHAR$

$\#(left \frown right) \leq maxsize$

$\#(left' \frown right') \leq maxsize$

$ch? = right_arrow$

$((right \neq \langle \rangle \wedge left' = left \frown \langle head\ right \rangle \wedge right' = tail\ right) \vee$
 $(right = \langle \rangle \wedge left' = left \wedge right' = right))$

Or

$T_Forward$

$\Delta Editor$

$ch? : CHAR$

$ch? = right_arrow$

$((right \neq \langle \rangle \wedge left' = left \frown \langle head\ right \rangle \wedge right' = tail\ right) \vee$
 $(right = \langle \rangle \wedge left' = left \wedge right' = right))$

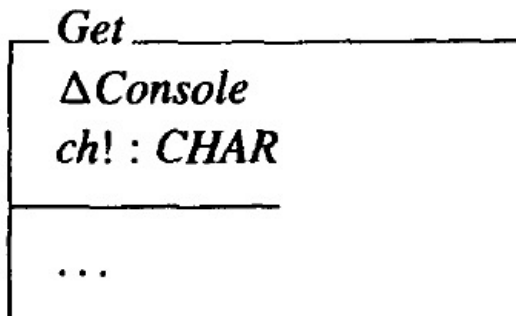


Other schema calculus operators

- We can define the ForwardTwo operation via using the schema composition operator

$$\textit{ForwardTwo} \hat{=} \textit{Forward} \circ \textit{Forward}$$

- We can use the schema piping for combining operations that communicate through input and output variables.
- For example, we wish to model the low-level interface to console keyboard from which our editor receives its input. The *Get* operation indicates the most recently pressed keyboard



Other schema calculus operators

- Then we could form a new operation G_Insert by piping the output of Get to the input of $Insert$:

$$G_Insert \hat{=} Get \gg Insert$$

- The pipe operator \gg ensures that the output and input variables $ch!$ and $ch?$ are merged

Schema types

We have defined four kinds of fundamental data types

- Free types, i.e., listing all members

$OP ::= plus \mid minus \mid times \mid divide$

- Basic types, declared as in $[X]$, instances are individuals
- Set types, declared as in $P X$, instances are sets
- Cartesian product types, declared as in $X * Y$, whose instances are tuples

Schema binding

- A binding is the formal realization of what we have been calling *a situation or a state*
- A binding is *an assignment of particular values to a collection of named variables*
- A binding resembles a tuple in that it is a composite object whose components can have different types.
- Components of a binding are distinguished by name.

An calendar example

- We try to define an appointment calendar
- The calendar includes the current day, month and year
- We need to define an operation the advances the date by one day
- First, we define data types

$DAY == 1 \dots 31$

$MONTH == 1 \dots 12$

$YEAR == \mathbb{Z}$

$DATE == DAY \times MONTH \times YEAR$

An calendar example

- We also need to know how many days in a particular month
 - Some months have 30 or 31 days
 - Feb has 28 days and 29 days in a leap year
 - We define the function *days*

$$\text{days} == \{1 \mapsto 31, 2 \mapsto 28, \dots, 12 \mapsto 31\}$$

- Then we can define the function *next* that takes a DATE and returns the next DATE

next : DATE \rightarrow DATE

$\forall d : \text{DAY}; m : \text{MONTH}; y : \text{YEAR} \bullet$

$(d < \text{days } m \wedge \text{next}(d, m, y) = (d + 1, m, y)) \vee$

$(d = \text{days } m \wedge m < 12 \wedge \text{next}(d, m, y) = (1, m + 1, y)) \vee$

$(d = \text{days } m \wedge m = 12 \wedge \text{next}(d, m, y) = (1, 1, y + 1))$

- Problems with this definition ?

An calendar example

- We also need to know how many days in a particular month
 - Some months have 30 or 31 days
 - Feb has 28 days and 29 days in a leap year
 - We define the function *days*

$$\text{days} == \{1 \mapsto 31, 2 \mapsto 28, \dots, 12 \mapsto 31\}$$

- Then we can define the function *next* that takes a DATE and returns the next DATE

next : DATE \rightarrow DATE

$\forall d : \text{DAY}; m : \text{MONTH}; y : \text{YEAR} \bullet$

$(d < \text{days } m \wedge \text{next}(d, m, y) = (d + 1, m, y)) \vee$

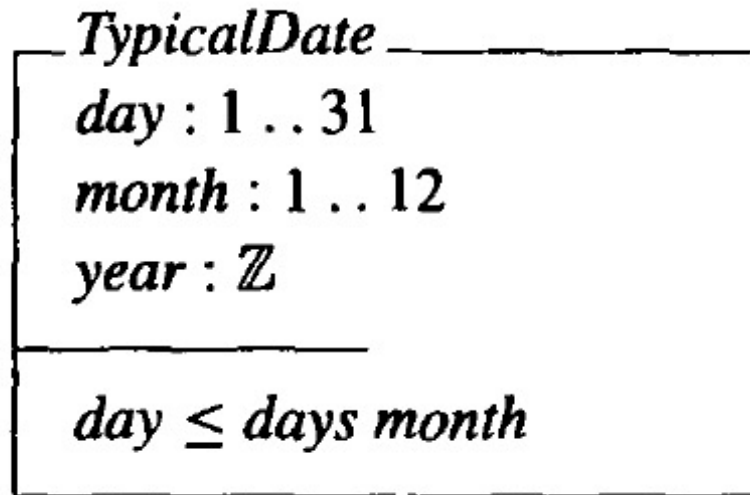
$(d = \text{days } m \wedge m < 12 \wedge \text{next}(d, m, y) = (1, m + 1, y)) \vee$

$(d = \text{days } m \wedge m = 12 \wedge \text{next}(d, m, y) = (1, 1, y + 1))$

- Problems with this definition ? (the leap year)

Calendar example

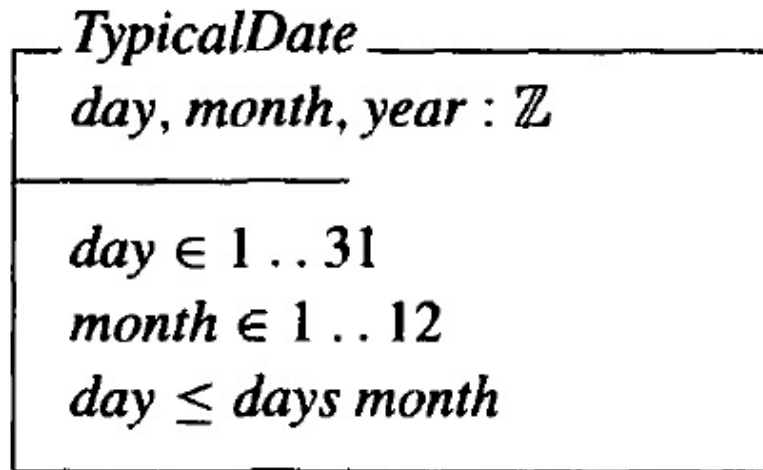
- Then we can define a TypicalDate schema



- However, it is not a normalized schema because we just specify the range but not the type of variables

Calendar example

- Then we can define a normalized TypicalDate schema



- However, the TypicalDate schema does not cover the leap year, i.e., there is an extra day, February 29th.
- Define another schema to cover the exceptional case, then use schema disjunction to combine the cases

Calendar example

- We define a prefix unary relation leap_ , so that “leap year” is true when year is a leap year
- Our “ leap_ ” is just a set of years: leap years occur every four years, excluding only centuries not divisible by 400 (2000 is a leap year, but not 1900)

$$(\text{leap_}) == \{ y : \mathbb{Z} \bullet 4 * y \} \setminus ((\{ y : \mathbb{Z} \bullet 100 * y \} \setminus \{ y : \mathbb{Z} \bullet 400 * y \}))$$

<i>Feb29</i>
<i>month, day, year : \mathbb{Z}</i>
<i>month = 2</i>
<i>day = 29</i>
<i>leap year</i>

$$\text{Date} \hat{=} \text{TypicalDate} \vee \text{Feb29}$$

Calendar example

- So the normalized Date Schema is:

Date

day, month, year : \mathbb{Z}

*$(day \in 1 \dots 31 \wedge day \leq days\ month \wedge month \in 1 \dots 12) \vee$
 $(day = 29 \wedge month = 2 \wedge leap\ year)$*



Calendar example

- Let's define the *weekday function* which tells use the day of the week for any date
- For example, July 20, 1967 was a Sunday
- We represent the days with numbers:
 - Sunday is 0, Monday is 1, and so on

weekday : *Date* \rightarrow 0 .. 6

$\forall d : \text{Date} \bullet \text{weekday}(d) = (d.\text{day} + d.\text{year} + d.\text{year} \text{ div } 4 + \dots) \bmod 7$

- Problem !
- Date is a whole set, but the argument of weekday is a single binding

Calendar example

- To solve this problem, Z provides the binding formation operator θ (the Greek letter theta).
- θ Date refers to the member of Date that is currently in scope

$$\begin{array}{|l} \text{weekday} : \text{Date} \rightarrow 0 \dots 6 \\ \hline \forall \text{Date} \bullet \text{weekday}(\theta \text{Date}) = (\text{day} + \text{year} + \text{year} \text{ div } 4 + \dots) \bmod 7 \end{array}$$

- θ operator can also be used in predicates to avoid some details

$$\begin{array}{|l} \exists \text{Editor} \text{ } \text{ } \\ \Delta \text{Editor} \\ \hline \theta \text{Editor}' = \theta \text{Editor}_1 \end{array}$$

Generic definitions

- Z also provides generic constructs that enable you to write definitions that apply to any type
- For example, we used the concatenation operator $\hat{}$ to join texts together: $\text{left}' = \text{left} \hat{} \langle \text{head right} \rangle$
- We can define a generic concatenation

$$\boxed{\begin{array}{l} [X] \\ \hline _ \hat{} _ : \text{seq } X \times \text{seq } X \rightarrow \text{seq } X \\ \hline \dots \end{array}}$$

- X is a formal generic parameter that stands for any type.

Generic schema

- Z also provides generic schemas
- We define a generic schema *Pool* which contains the generic parameter *RESOURCE*
- *RESOURCE* can be used to represent any resource, such as memory pages, disk blocks, or processors

Pool [*RESOURCE*]

owner : *RESOURCE* \rightarrow *USER*

free : \mathbb{P} *RESOURCE*

$(\text{dom } \textit{owner}) \cup \textit{free} = \textit{RESOURCE}$

$(\text{dom } \textit{owner}) \cap \textit{free} = \emptyset$



Formal reasoning

So far, we introduced how to use Z to define types and schemas. However, how do we know the schemas and operations are correct and can represent the user's requirements without a missing.

We can use formal reasoning to validate a mathematical model against requirements

A model is called valid if its properties satisfy the intent of the requirements from users.

Formal reasoning can also show how one model is related to another, and can verify that code implements its specification.

In Z, an exercise in formal reasoning is also called a proof.

Calculation and proof

- Formal reasoning means reasoning with formulas.
- The arithmetic calculations are examples of formal reasoning.
- For example, “A train moves at a constant velocity of sixty km per hour. How far does the train travel in four hours?”
- To solve this problem we can express the problem in Z

distance, velocity, time : \mathbb{N}

distance = *velocity* * *time*

velocity = 60

time = 4



Calculation and proof

- A formal proof of the distance = 240 km can be:

$distance = velocity * time$	[Definition]
$= 60 * time$	[velocity = 60]
$= 60 * 4$	[time = 4]
$= 240$	[Arithmetic]

- If we change the problem to say velocity < 60 km/hour, then the proof becomes:

$distance = velocity * time$	[Definition]
$< 60 * time$	[velocity < 60]
$= 60 * 4$	[time = 4]
$= 240$	[Arithmetic]

Calculation and proof

- Calculations need not be arithmetic
- Set operators can also be used in the proof

philip : PERSON

adhesives, materials, research, manufacturing : \mathbb{P} PERSON

adhesives \subseteq materials

materials \subseteq research

philip \in adhesives

- Can we proof Philip works in the research division?

philip \in adhesives

[Definition]

\subseteq materials

[Definition]

\subseteq research

[Definition]

Calculation and proof

- Another example to find the value of x , given $2x + 7 = 13$

$$\begin{array}{|l} x : \mathbb{Z} \\ \hline 2 * x + 7 = 13 \end{array}$$

We simply solve for x

$$2 * x + 7 = 13 \quad \text{[Definition.]}$$

$$\Leftrightarrow 2 * x = 13 - 7 \quad \text{[Subtract 7 from both sides.]}$$

$$\Leftrightarrow 2 * x = 6 \quad \text{[Arithmetic.]}$$

$$\Rightarrow (2 * x) \text{ div } 2 = 6 \text{ div } 2 \quad \text{[Divide both sides by 2.]}$$

$$\Leftrightarrow x = 6 \text{ div } 2 \quad \text{[Division on left side, algebra]}$$

$$\Leftrightarrow x = 3 \quad \text{[Division on right side, arithmetic]}$$

This completes our proof of the predicate $2 * x + 7 = 13 \Rightarrow x = 3$.

Laws

- A proof is valid only when every step is justified, but we need an authority to confirm that every step is justified.
- In Z, the authority, i.e., a collection of formulas, is called laws. Laws are NOT variables, but are place-holders.
- A variable always denotes a particular value, but a place-holder represents any expression of the appropriate type.
- Also, a predicate may be true or false depending on the values of its variables, but a law is always true.
- Predicates express facts that are particular to some specific situation, but laws express rules that are universally applicable.

Laws

- For example, $2 * n = 6$ is a predicate but not a law
 - It is only true when $n = 3$
- However, $0 * n = 0$ is a law because it is always true for all n .
- Another example, $n = d * q + r$ is a predicate
 - It is true when $n = 7$, $q = 2$, $d = 3$ and $r = 1$, and some other cases. However it is false when $n = 7$, $q = 2$, $d = 3$ and $r = 0$
- $n = d * (n \text{ div } d) + (n \text{ mod } d)$ for all d not equals 0
 - It is a law and always true

Checking specifications

- We can use formal reasoning to check our work for certain kinds of errors and oversights
- This is one of the most important qualities that distinguishes a formal method from informal ones
- For example, for the simple editor system, we can say the editor must have an initial state, i.e., $\exists \textit{State} \bullet \textit{Init}$ must be true.

$$\exists \textit{left}, \textit{right} : \textit{TEXT} \mid \#(\textit{left} \frown \textit{right}) \leq \textit{maxsize} \bullet \textit{left} = \textit{right} = \langle \rangle$$

Precondition calculation

- Many actual program fails because programmers did not account for all the preconditions
- For example, in the simple editor, our first attempt to define the operation that moves the cursor forward was inadequate because it failed to account for the situation where the cursor is already at the end of the file.
- If we had implemented that first version, we might have produced a faulty program that could crash and losing all the user's work
- We can calculate the precondition of any operation defined by a Z schema

Precondition calculation

- The precondition indicates unprimed “before” variables and input variables only

<i>Forward</i>
$\Delta Editor$
$ch? : CHAR$
$ch? = right_arrow$
$right \neq \langle \rangle$
$left' = left \frown \langle head(right) \rangle$
$right' = tail(right)$

- The precondition is $ch? = right_arrow \wedge right \neq \langle \rangle$
- An operation is called total when the precondition covers all possibilities. Clearly the above precondition is not total because it doesn't cover the case when $right = \langle \rangle$

Precondition calculation

- So we must define $T_Forward$ operation to account for that case

$T_Forward$
$\Delta Editor$
$ch? : CHAR$
$ch? = right_arrow$
$((right \neq \langle \rangle \wedge left' = left \wedge \langle head\ right \rangle \wedge right' = tail\ right) \vee$ $(right = \langle \rangle \wedge left' = left \wedge right' = right))$

- The precondition is $(right \neq \langle \rangle) \vee (right = \langle \rangle)$
- This precondition is total

Precondition calculation

- Sometime the precondition of an operation is implicit, but we can calculate the precondition with Z schema

<i>Insert</i>
<i>ΔEditor</i>
<i>ch? : CHAR</i>
<i>ch? ∈ printing</i>
<i>left' = left ∪ {ch?}</i>
<i>right' = right</i>

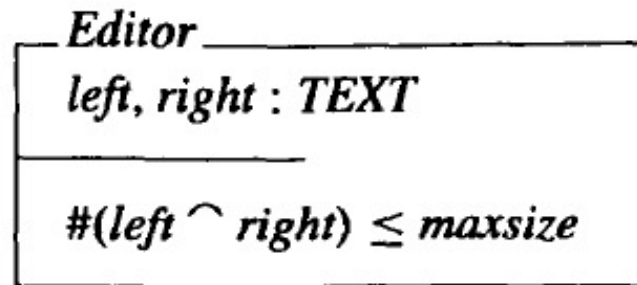
- Does this operation always work?
- Does it have some preconditions that aren't obvious?
- Let's calculate the precondition

Precondition calculation

- We know that there must exist a state of Editor that satisfies the predicate of the insert operation schema, i.e.,

$$\exists \textit{Editor}' \bullet \textit{Insert}$$

- Because



- So

$$\begin{aligned} \exists \textit{left}', \textit{right}' : \textit{TEXT} \mid \#(\textit{left}' \frown \textit{right}') \leq \textit{maxsize} \bullet \\ \textit{ch?} \in \textit{printing} \wedge \textit{left}' = \textit{left} \frown \{\textit{ch?}\} \wedge \textit{right}' = \textit{right} \end{aligned}$$



Precondition calculation

- It equals

$\exists left', right' : TEXT \bullet$

$$ch? \in printing \wedge \#(left' \frown right') \leq maxsize \wedge \\ left' = left \frown \langle ch? \rangle \wedge right' = right$$

- We can combine the left' and right' variables, then we obtain

$$\#ch? \in printing \wedge ((left \frown \langle ch? \rangle) \frown right) \leq maxsize$$

- The standard proof for this precondition can be as follows.

Precondition calculation

$\exists Editor' \bullet Insert$

[Definition of precondition]

$\Leftrightarrow \exists left', right' : TEXT \mid \dots \bullet \dots$

[Expand schemas]

$\Leftrightarrow \exists left', right' : TEXT \bullet \dots \wedge \dots$

[Restricted \exists -quantifier]

$\Leftrightarrow pr \wedge \#((left \hat{\ } \langle ch? \rangle) \hat{\ } right) \leq maxsize$

[One-point rule]

$\Leftrightarrow pr \wedge \#left + \#\langle ch? \rangle + \#right \leq maxsize$

$[\#(s \hat{\ } t) = \#s + \#t]$

$\Leftrightarrow pr \wedge \#left + 1 + \#right \leq maxsize$

$[\#\langle x \rangle = 1]$

$\Leftrightarrow pr \wedge \#left + \#right < maxsize$

[Arithmetic]

- This example is obvious, but many software failures results from errors that seems obvious in retrospect
- In more complicated applications, the calculation sometimes reveals unexpected preconditions