Notation Supporting Grammars

> ¬ Real programmers do not use Greek letters.
> *Anonymous*

Conventions

The theory of compilers can be expressed in terms of logic, sets, and relations. When a symbol or formula is defined in terms of previously established material, the form

$$new \stackrel{\text{def}}{=} old$$

is used. Otherwise '=' means equality for whatever type of operands it has. Sometimes a definition is not complete without further constraints on its components. In this case the constraint is also decorated with $^{\text{def}}$. For example

$$e1 \stackrel{\text{def}}{\subseteq} e2$$

means that $e1$ is constrained to be a subset of $e2$ in addition to whatever information has been previously given.

Conventions for operator binding reveal the meaning of an expression without a blizzard of parentheses. We know that

$$a + b/c$$

means $a + (b/c)$, not $(a + b)/c$ because we have been taught that division binds more tightly than addition. Conventions such as this are used both in mathematical notation and in programming languages. Arithmetic operators in the mathematical notation used here are presumed to bind more tightly than all others. When operators of equal strength are used, operands group left-to-right.

The use of binding strength on these pages is informal; the purpose is to ease the task for an intelligent reader rather than to enable formal analysis.

Do not confuse the definitions below with axiom systems or any other kind of mathematics within which there is an expectation of formal derivations. These are descriptions, useful for communication, but not intended for proofs.

Typography

Fonts are used to imply the class of thing being represented. In particular, computer input is monowidth, variables are lower case italic, sets are upper case italic. Sequences are lower case Greek letters. Relations are upper case Greek letters. Caligraphic script is used for higher level constructs. The following glyphs are used for Greek.

α	A	alpha	ι	I	iota	ρ	P	rho
β	B	beta	κ	K	kappa	σ	Σ	sigma
γ	Γ	gamma	λ	Λ	lambda	τ	T	tau
δ	Δ	delta	μ	M	mu	υ	Υ	upsilon
ϵ	E	epsilon	ν	N	nu	ϕ	Φ	phi
ζ	Z	zeta	ξ	Ξ	xi	χ	X	chi
η	H	eta	o	O	omicron	ψ	Ψ	psi
θ	Θ	theta	π	Π	pi	ω	Ω	omega

Propositions and Predicates

The operators for the propositional calculus are given, in order of binding strength, in Table 1. Logical negation is most tightly bound.

$\Leftrightarrow$	logical equivalence
$\Rightarrow$	logical implication
$\vee$	logical or
$\wedge$	logical and
$\neg$	logical negation

Table 1: Propositional Operators

One can write formulas such as the following tautology

$$\neg(p \vee q) \Leftrightarrow (\neg p \wedge \neg q)$$

Table 2: DeMorgan's Law

The extension of the propositional calculus to the predicate calculus introduces the two quantifiers '$\forall$' (read as 'for all') and '$\exists$' (read as 'there exists'). They are like declarations in programming languages in that they introduce new variables of a specific type for part of a logical expression. The form of quantification is given by example in Table 3. The scope of the bound variables is the predicate to the right of the separating dot. Parentheses are used if the scope

of operators or quantifiers is unclear. If the universe of quantification is obvious from the context, it is implicit. Otherwise an explicit set membership will be included as part of the predicate. A variable bound by no quantifier is called *free*. Such a variable is implicitly universally quantified.

$$\forall x.\ Prime(x) \Leftrightarrow 1 < x \land \neg \exists y.\ 1 < y \land y < x \land (x/y) * y = x$$

Table 3: Example Predicate

Sets

Sets consist of elements taken from some universe. The universe might be a specific finite set, or it might be countably infinite like the integers. Knowing the universe is often part of understanding a formula in set theory. Let A and B be sets and P be a predicate.

$\{\}$	empty set
$\overline{\{\}}$	universe
$\{1, 2, 3\}$	explicit finite set
$a \in A$	set membership, a is in A
$\{x \mid P(x)\}$	set of x such that $P(x)$
$\text{choice}(A)$	nondeterministic selection
$\text{size}(A)$	size of set A

Table 4: Set Primitives

$A = B$	$\stackrel{\text{def}}{=} x \in A \Leftrightarrow x \in B$	set equivalence
$A \subseteq B$	$\stackrel{\text{def}}{=} x \in A \Rightarrow x \in B$	subset
$A \subset B$	$\stackrel{\text{def}}{=} A \subseteq B \land B \nsubseteq A$	proper subset
$A \cup B$	$\stackrel{\text{def}}{=} \{x \mid x \in A \lor x \in B\}$	union
$A \cap B$	$\stackrel{\text{def}}{=} \{x \mid x \in A \land x \in B\}$	intersection
$A - B$	$\stackrel{\text{def}}{=} \{x \mid x \in A \land x \notin B\}$	difference
$\overline{A}$	$\stackrel{\text{def}}{=} \{x \mid x \in \overline{\{\}} \land x \notin A\}$	complement w.r.t. universe
2^A	$\stackrel{\text{def}}{=} \{B \mid B \subseteq A\}$	powerset

Table 5: Set Operators

The operators above are presented in relative binding strength. The set operators bind more strongly than logical operators, and less strongly than arithmetic operators. The function 'choice' is undefined if it is applied to an empty set; it is some value from the set otherwise.

Exercises

Each of the following formulas is either true or false. If it is false, find a modification that makes it true.

1. $\text{size}(2^A) = 2^{\text{size}(A)}$

2. $((A - B) \cap B) = \{\}$

3. $\text{choice}(\overline{\{\}}) \in \overline{\{\}}$

4. $\text{choice}(A) \in A$

5. $\text{size}(A - \{\text{choice}(A)\}) = \text{size}(A) - 1$

Ordered Pairs and Tuples

One of the most useful concepts is that of *pair*. It allows us to express associations of many kinds. There is an obvious generalization to *tuple* which is also sometimes of use.

Let C and D be sets, c and d be elements:

$$\langle c, d \rangle \qquad\qquad\qquad\qquad\qquad \text{ordered pair}$$
$$C \times D \quad \overset{\text{def}}{=} \{\langle c, d \rangle \mid c \in C \wedge d \in D\} \quad \text{cross product}$$

Table 6: Ordered Pairs

Sequences

Nearly everything in computing comes in sequences. Partly it is in the nature of machinery moving in time, time providing the sequencing. Sometimes it in the nature of space, with "next to" providing the sequencing.

Unless otherwise indicated, lower case Greek letters denote sequences. Let d be an element (a sequence of length one), α, β be sequences, A and B be sets of sequences.

λ		empty sequence
$\alpha\beta$		catenation of sequences
$\alpha\lambda = \lambda\alpha$	$\overset{\text{def}}{=} \alpha$	catenation of λ
$A \circ B$	$\overset{\text{def}}{=} \{\alpha\beta \mid \alpha \in A \wedge \beta \in B\}$	catenation of sets of sequences
$\text{length}(\lambda)$	$\overset{\text{def}}{=} 0$	length of empty sequence
$\text{length}(\alpha\beta)$	$\overset{\text{def}}{=} \text{length}(\alpha) + \text{length}(\beta)$	length of sequence
α_n	$\overset{\text{def}}{=} \text{d} \mid \alpha = \beta\text{d}\gamma \wedge \text{length}(\beta\text{d}) = n$	subscript selection
A^0	$\overset{\text{def}}{=} \{\lambda\}$	sequences of length 0
A^n	$\overset{\text{def}}{=} A \circ A^{n-1}$	sequences of positive length n
A^*	$\overset{\text{def}}{=} \bigcup_{i=0}^{\infty} A^i$	all finite sequences
A^+	$\overset{\text{def}}{=} \bigcup_{i=1}^{\infty} A^i$	non-empty finite sequences
$A^{1/*}$	$\overset{\text{def}}{=} \{\text{d} \mid \beta\text{d}\gamma \in A\}$	elements in sequences in A

Table 7: Notation for Sequences

The notation A^* means all finite sequences of elements taken from A. The inverse, $B^{1/*}$, is the elements which make up the sequences in B. The importance of sequences is that we regard a language as a set of sequences of words. If A is a dictionary, then A^* is everything you can ever say (meaningful or not).

Exercises

6. Is $\text{size}(A \times B) = \text{size}(A) * \text{size}(B)$?

7. Is $A \subseteq (A^*)^{1/*}$?

8. What is $\{\}^*$?

9. What is $\{\}^+$?

10. What is $\text{size}(\{\text{d}\}^7)$?

Relations

Relations are an abstraction applied both to pairs and to sequences. Relations have a useful calculus of their own. The answer we seek is often expressed as a relation. A relation is a set of ordered pairs.

Let A be a set, Φ and Ψ be relations.

$\mathcal{D}(\Phi)$	$\stackrel{\text{def}}{=} \{x \mid \langle x, y \rangle \in \Phi\}$	domain
$\mathcal{R}(\Phi)$	$\stackrel{\text{def}}{=} \{y \mid \langle x, y \rangle \in \Phi\}$	range
$A \triangleleft \Phi$	$\stackrel{\text{def}}{=} \{\langle x, y \rangle \mid \langle x, y \rangle \in \Phi \wedge x \in A\}$	domain restriction
$\Phi \triangleright A$	$\stackrel{\text{def}}{=} \{\langle x, y \rangle \mid \langle x, y \rangle \in \Phi \wedge y \in A\}$	range restriction
$\Psi \circ \Phi$	$\stackrel{\text{def}}{=} \{\langle x, z \rangle \mid \exists y.\ \langle x, y \rangle \in \Psi \wedge \langle y, z \rangle \in \Phi\}$	composition
Φ^0	$\stackrel{\text{def}}{=} \{\langle x, x \rangle \mid x \in \overline{\{\}}\}$	identity relation
Φ^{-1}	$\stackrel{\text{def}}{=} \{\langle y, x \rangle \mid \langle x, y \rangle \in \Phi\}$	inverse relation
Φ^n	$\stackrel{\text{def}}{=} \Phi \circ \Phi^{n-1}$	transitive extension $(n > 0)$
Φ^*	$\stackrel{\text{def}}{=} \bigcup_{i=0}^{\infty} \Phi^i$	zero or more steps
Φ^+	$\stackrel{\text{def}}{=} \bigcup_{i=1}^{\infty} \Phi^i$	one or more steps
$\Phi(x)$	$\stackrel{\text{def}}{=} \text{choice}(\mathcal{R}(\{x\} \triangleleft \Phi))$	function application

Table 8: Relations

When a relation is also a function, application is deterministic. The superscript '$*$' has two reasonable interpretations. Since any relation is also a set, '$*$' could be used to signify a sequence of pairs. In fact it is used instead for the reflexive transitive completion of the relation; a matter of iterating on composition.

Exercises

11. Is $A \triangleleft \Phi = (A \times \overline{\{\}}) \cap \Phi$?

12. Is $\Phi \triangleright A = \Phi \cap (\overline{\{\}} \times A)$?

13. Let P be the universe of living people, M be all men, W be all women. $P = M \cup W$. Let $Married \subseteq M \times W$.[1]

    (a) What are $\mathcal{R}(Married)$ and $\mathcal{D}(Married)$?

    (b) What do you call x where $\text{size}(\mathcal{R}(\{x\} \triangleleft Married)) > 1$?

    (c) What do you call $M - \mathcal{D}(Married)$?

    (d) Suppose $y \in W$. What is $Married^{-1}(y)$?

14. Continuing the previous exercise, let $Mother \subseteq P \times W$ and $Father \subseteq P \times M$.

    (a) Let $c = \text{choice}(P)$. When is $Mother(c)$ defined?

    (b) What is $Mother^{-1} \cup Father^{-1}$?

    (c) What is $\overline{\mathcal{D}(Mother \cup Father)}$?

    (d) What is $Mother \cap Father$?

    (e) What is $Mother^{-1}$?

    (f) What is $Mother^2$?

    (g) What is $Mother^*$?

Ordered Sets

Suppose R is a relation on $A \times A$. R is a an ordering if it antisymmetric and transitive. For convenience, we use the infix operator $<$ to describe orderings in general. There are two cases of interest.

Partial Order

$\forall a\ b \in A.\ \neg(a < b \wedge b < a)$
$\forall a\ b\ c \in A.\ a < b \wedge b < c \Rightarrow a < c$

Simple Order

A simple order is a partial order where in addition

$\forall a\ b \in A.\ a < b \vee a = b \vee b < a$
$\min(A) \stackrel{\text{def}}{=} \mathcal{D}(R) - \mathcal{R}(R)$
$\max(A) \stackrel{\text{def}}{=} \mathcal{R}(R) - \mathcal{D}(R)$

---

[1]If current legalisms get in your way for this exercise, understand the relation $Married$ in the traditional sense.

Exercises

15. Suppose P, M and W refer to the living and the dead. Describe Adam and Eve.