

# Specification of Source §2 Typed—2022 edition

Martin Henz, Jingjing Zhao

National University of Singapore  
School of Computing

January 18, 2023

The language Source is the official language of the textbook *Structure and Interpretation of Computer Programs, JavaScript Adaptation*. Source is a sublanguage of ECMAScript 2018 (9<sup>th</sup> Edition) and defined in the documents titled “Source §*x*”, where *x* refers to the respective textbook chapter.

Source §2 Typed is a variant of Source §2 that introduces type syntax and type checking.

## 1 Syntax

A Source program is a *program*, defined using Backus-Naur Form<sup>1</sup> as follows:

---

<sup>1</sup> We adopt Henry Ledgard’s BNF variant that he described in *A human engineered variant of BNF*, ACM SIGPLAN Notices, Volume 15 Issue 10, October 1980, Pages 57-62. In our grammars, we use **bold** font for keywords, *italics* for syntactic variables,  $\epsilon$  for nothing,  $x \mid y$  for *x* or *y*,  $[ x ]$  for an optional *x*,  $x \dots$  for zero or more repetitions of *x*, and  $( x )$  for clarifying the structure of BNF expressions.

<i>program</i>	::= <i>import-directive... statement...</i>	<b>program</b>
<i>import-directive</i>	::= <b>import</b> { <i>import-names</i> } <b>from</b> <i>string</i> ;	import directive
<i>import-names</i>	::= $\epsilon$   <i>import-name</i> ( , <i>import-name</i> )...	import name list
<i>import-name</i>	::= <i>name</i>   <i>name as name</i>	import name
<i>statement</i>	::= <b>const</b> <i>name</i> [ : <i>type</i> ] = <i>expression</i> ;   <b>function</b> <i>name</i> ( <i>names</i> ) [ : <i>type</i> ] <i>block</i>   <b>type</b> <i>name</i> = <i>type</i> ;   <b>return</b> <i>expression</i> ;   <i>if-statement</i>   <i>block</i>   <i>expression</i> ;   <b>debugger</b> ;	constant declaration function declaration type alias declaration return statement conditional statement block statement expression statement breakpoint
<i>names</i>	::= $\epsilon$   <i>name</i> [ : <i>type</i> ] ( , <i>name</i> [ : <i>type</i> ] )...	name list
<i>if-statement</i>	::= <b>if</b> ( <i>expression</i> ) <i>block</i>   <b>else</b> ( <i>block</i>   <i>if-statement</i> )	conditional statement
<i>block</i>	::= { <i>statement...</i> }	block statement
<i>expression</i>	::= <i>number</i>   <b>true</b>   <b>false</b>   <i>string</i>   <b>null</b>   <i>name</i>   <i>expression</i> <i>binary-operator</i> <i>expression</i>   <i>unary-operator</i> <i>expression</i>   <i>expression</i> <i>binary-logical</i> <i>expression</i>   <i>expression</i> ( <i>expressions</i> )   ( <i>names</i> ) => <i>expression</i>   ( <i>names</i> ) => <i>block</i>   <i>expression</i> ? <i>expression</i> : <i>expression</i>   ( <i>expression</i> )   <i>expression as type</i>	primitive number expression primitive boolean expression primitive string expression primitive list expression name expression binary operator combination unary operator combination logical composition function application lambda expression (expr. body) lambda expression (block body) conditional expression parenthesised expression as expression
<i>binary-operator</i>	::= +   -   *   /   %   ===   !==   >   <   >=   <=	binary operator
<i>unary-operator</i>	::= !   -	unary operator
<i>binary-logical</i>	::= &&	logical composition symbol
<i>expressions</i>	::= $\epsilon$   <i>expression</i> ( , <i>expression</i> )...	argument expressions
<i>type</i>	::= <i>number</i>   <i>boolean</i>   <i>string</i>   <i>undefined</i>   <i>null</i>   <i>void</i>   <i>any</i>   <i>number</i>   <i>string</i>   <b>true</b>   <b>false</b>   <i>name</i>   ( <i>names</i> ) => <i>type</i>   <i>type</i>   <i>type</i>   <i>Pair</i> < <i>type</i> , <i>type</i> >   <i>List</i> < <i>type</i> >	basic type literal type type reference function type union type pair type list type

## Restrictions

- Return statements are only allowed in bodies of functions.
- There cannot be any newline character between `return` and *expression* in return statements.<sup>2</sup>
- There cannot be any newline character between `( name | ( parameters ) )` and `=>` in function definition expressions.<sup>3</sup>
- Implementations of Source are allowed to treat function declaration as **syntactic sugar for constant declaration**.<sup>4</sup> Source programmers need to make sure that functions are not called before their corresponding function declaration is evaluated.

## Import directives

Import directives allow programs to import values from modules and bind them to names, whose scope is the entire program in which the import directive occurs. Import directives can only appear at the top-level. All names that appear in import directives must be distinct, and must also be distinct from all top-level variables. The Source specifications do not specify how modules are programmed.

## Logical Composition

### Conjunction

$$expression_1 \ \&\& \ expression_2$$

stands for

$$expression_1 \ ? \ expression_2 \ : \ \mathbf{false}$$

### Disjunction

$$expression_1 \ || \ expression_2$$

stands for

$$expression_1 \ ? \ \mathbf{true} \ : \ expression_2$$

## Names

Names<sup>5</sup> start with `_`, `$` or a letter<sup>6</sup> and contain only `_`, `$`, letters or digits<sup>7</sup>. Restricted words<sup>8</sup> are not allowed as names.

Valid names are `x`, `_45`, `$$` and `π`, but always keep in mind that programming is communicating and that the familiarity of the audience with the characters used in names is an important aspect of program readability.

## Numbers

We use decimal notation for numbers, with an optional decimal dot. “Scientific notation” (multiplying the number with  $10^x$ ) is indicated with the letter `e`, followed by the exponent `x`. Examples for numbers are `5432`, `-5432.109`, and `-43.21e-45`.

<sup>2</sup> Source inherits this syntactic quirk of JavaScript.

<sup>3</sup>ditto

<sup>4</sup>ECMAScript prescribes “hoisting” of function declarations to the beginning of the surrounding block. Programs that rely on this feature will run fine in JavaScript but might encounter a runtime error “Cannot access name before initialization” in a Source implementation.

<sup>5</sup> In [ECMAScript 2020 \(9th Edition\)](#), these names are called *identifiers*.

<sup>6</sup> By *letter* we mean [Unicode](#) letters (L) or letter numbers (NI).

<sup>7</sup> By *digit* we mean characters in the [Unicode](#) categories Nd (including the decimal digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9), Mn, Mc and Pc.

<sup>8</sup> By *restricted word* we mean any of: `arguments`, `await`, `break`, `case`, `catch`, `class`, `const`, `continue`, `debugger`, `default`, `delete`, `do`, `else`, `enum`, `eval`, `export`, `extends`, `false`, `finally`, `for`, `function`, `if`, `implements`, `import`, `in`, `instanceof`, `interface`, `let`, `new`, `null`, `package`, `private`, `protected`, `public`, `return`, `static`, `super`, `switch`, `this`, `throw`, `true`, `try`, `typeof`, `var`, `void`, `while`, `with`, `yield`. These are all words that cannot be used without restrictions as names in the strict mode of ECMAScript 2020.

## Strings

Strings are of the form "*double-quote-characters*", where *double-quote-characters* is a possibly empty sequence of characters without the character " and without the newline character, of the form '*single-quote-characters*', where *single-quote-characters* is a possibly empty sequence of characters without the character ' and without the newline character, and of the form '*backquote-characters*', where *backquote-characters* is a possibly empty sequence of characters without the character \. Note that newline characters are allowed as *backquote-characters*.

The following characters can be represented in strings as given:

- horizontal tab: `\t`
- vertical tab: `\v`
- nul char: `\0`
- backspace: `\b`
- form feed: `\f`
- newline: `\n`
- carriage return: `\r`
- single quote: `\'`
- double quote: `\"`
- backslash: `\\`

Unicode characters can be used in strings using `\u` followed by the hexadecimal representation of the unicode character, for example `'\uD83D\uDC04'`.

## Comments

In Source, any sequence of characters between `/*` and the next `*/` is ignored. After `/**` any characters until the next newline character is ignored.

## 2 Type System

In Source §2 Typed, the Source §2 syntax is expanded to include type syntax such as type annotations and type aliases. This allows names to be explicitly typed, and for type checks to be performed.

Support for `typeof` operations is also added to Source §2 Typed.

### 2.1 Type Environment

In order to keep track of the type of names in a program, we define a *type environment*, denoted by  $\Gamma$ . More formally, the partial function  $\Gamma$  from names to types expresses a context, in which a name  $x$  is associated with type  $\Gamma(x)$ .

We define a relation  $\Gamma[x \leftarrow t]\Gamma'$  on type environments  $\Gamma$ , names  $x$ , types  $t$ , and type environments  $\Gamma'$ , which constructs a type environment that behaves like the given one, except that the type of  $x$  is  $t$ . More formally, if  $\Gamma[x \leftarrow t]\Gamma'$ , then  $\Gamma'(y)$  is  $t$ , if  $y = x$  and  $\Gamma(y)$  otherwise. Obviously, this uniquely identifies  $\Gamma'$  for a given  $\Gamma$ ,  $x$ , and  $t$ , and thus the type environment extension relation is functional in its first three arguments.

The set of names, on which a type environment  $\Gamma$  is defined, is called the domain of  $\Gamma$ , denoted by  $dom(\Gamma)$ .

For each non-overloaded primitive operator, we add a binding to our initial type environment  $\Gamma_0$  as follows:

```

∅[-₂ ← (number, number) → number]
[* ← (number, number) → number]
[/ ← (number, number) → number]
[% ← (number, number) → number]
[&& ← (boolean, T) → boolean | T]
[|| ← (boolean, T) → boolean | T]
[! ← boolean → boolean]
[-₁ ← number → number]
[typeof ← any → string]Γ₋₂

```

The overloaded binary primitives (with the exception of +, the handling of which will be elaborated in **Typing Relations**) are handled as follows:

```

Γ₋₂[=== ← (string, string) → boolean | (number, number) → boolean]
[!== ← (string, string) → boolean | (number, number) → boolean]
[> ← (string, string) → boolean | (number, number) → boolean]
[>= ← (string, string) → boolean | (number, number) → boolean]
[< ← (string, string) → boolean | (number, number) → boolean]
[<= ← (string, string) → boolean | (number, number) → boolean]Γ₋₁

```

The Source §2 standard library functions and constants have their types defined as follows:

```

Γ₋₁ [ display      ← any ]
    [ error      ← any ]
    [ Infinity   ← number ]
    [ is_boolean ← any      → boolean ]
    [ is_function ← any      → boolean ]
    [ is_number  ← any      → boolean ]
    [ is_string  ← any      → boolean ]
    [ is_undefined ← any    → boolean ]
    [ math_abs   ← number   → number ]
    [ math_acos  ← number   → number ]
    [ math_acosh ← number   → number ]
    [ math_asin  ← number   → number ]
    [ math_asinh ← number   → number ]
    [ math_atan  ← number   → number ]
    [ math_atan2 ← (number, number) → number ]
    [ math_atanh ← number   → number ]
    [ math_cbrt  ← number   → number ]
    [ math_ceil  ← number   → number ]
    [ math_clz32 ← number   → number ]
    [ math_cos   ← number   → number ]
    [ math_cosh  ← number   → number ]
    [ math_exp   ← number   → number ]
    [ math_expm1 ← number   → number ]
    [ math_floor ← number   → number ]
    [ math_fround ← number   → number ]
    [ math_hypot ← any      → number ]
    [ math_imul  ← (number, number) → number ]
    [ math_LN2   ← number   → number ]
    [ math_LN10  ← number   → number ]
    [ math_log   ← number   → number ]
    [ math_log1p ← number   → number ]
    [ math_log2  ← number   → number ]

```

[	math_LOG2E	←	number		]
[	math_log10	←	number	→	number
[	math_LOG10E	←	number		]
[	math_max	←	any		]
[	math_min	←	any		]
[	math_PI	←	number		]
[	math_pow	←	(number, number)	→	number
[	math_random	←	()	→	number
[	math_round	←	number	→	number
[	math_sign	←	number	→	number
[	math_sin	←	number	→	number
[	math_sinh	←	number	→	number
[	math_sqrt	←	number	→	number
[	math_SQRT1_2	←	number		]
[	math_SQRT2	←	number		]
[	math_tan	←	number	→	number
[	math_tanh	←	number	→	number
[	math_trunc	←	number	→	number
[	NaN	←	number		]
[	parse_int	←	(string, number)	→	number
[	prompt	←	string	→	string
[	get_time	←	()	→	number
[	stringify	←	any	→	string
[	undefined	←	undefined		]
[	null	←	null		]
[	pair	←	( $T_1, T_2$ )	→	$Pair< T_1, T_2 >$
[	head	←	$Pair< T_1, T_2 >$	→	$T_1$
[	tail	←	$Pair< T_1, T_2 >$	→	$T_2$
[	list	←	( $T_1, \dots, T_n$ )	→	$List< T_1   \dots   T_n >$
[	is_pair	←	any	→	boolean
[	is_null	←	any	→	boolean
[	is_list	←	any	→	boolean
					]
					$\Gamma_0$

## 2.2 Success Types

In order for type checks to be performed in Source §3 Typed, we introduce the notion of success types.

We first define the special `any` type:

**Definition 2.1** *any is the union of all possible types.*

Success typing in Source Typed is defined as follows:

**Definition 2.2** *Type  $t'$  is a success type of type  $t$  if  $\exists x(x \in t \wedge x \in t')$ . Alternatively:  $t \wedge t' \neq \emptyset$ .*

In Source Typed, type checks are performed by checking that the actual type is a success type of the expected type. This means that type errors will be thrown if and only if a definite clash in types at runtime is detected. Given that `any` is the union of all possible types, this also means that the `any` type is guaranteed not to produce any type errors.

## 2.3 Typing Relations

To perform type checking on the program, typing relations are applied to every statement and expression in the program.

Names that do not have a type declared will be assumed to have the `any` type.

### 2.3.1 Typing Relations on Expressions

The derived type of primitive expressions is their literal type, which is an element of its corresponding basic type.

$$\frac{}{\Gamma \vdash n : \textit{literal type } n} \quad \frac{}{\Gamma \vdash s : \textit{literal type } s}$$

where  $n$  denotes any literal number and  $s$  denotes any literal string.

$$\frac{}{\Gamma \vdash \mathbf{true} : \textit{literal type true}} \quad \frac{}{\Gamma \vdash \mathbf{false} : \textit{literal type false}}$$

For names, the type must be derived from the type environment.

$$\frac{}{\Gamma \vdash x : \Gamma(x)}$$

For function applications (including applications of binary and unary operators), the following two type rules are used, depending on the type of  $E_0$ .

$$\frac{\Gamma \vdash E_0 : (t_1, \dots, t_n) \rightarrow t \quad \Gamma \vdash E_1 : t'_1, \dots, \Gamma \vdash E_n : t'_n \quad (\forall 1 \leq i \leq n)(t'_i \wedge t_i \neq \emptyset)}{\Gamma \vdash (E_0 ( E_1, \dots, E_n )) : t}$$

$$\frac{\Gamma \vdash E_0 : \textit{any} \quad \Gamma \vdash E_1 : t'_1, \dots, \Gamma \vdash E_n : t'_n \quad (\forall 1 \leq i \leq n)(t'_i \wedge \textit{any} \neq \emptyset)}{\Gamma \vdash (E_0 ( E_1, \dots, E_n )) : \textit{any}}$$

The type of the operator must be a function type with the right number of parameters, and the type of every argument must be a success type of the corresponding parameter type of the function type. If all of the conditions are met, the type of the function application is the same as the return type of the function type that is the type of the operator. If the type of the operator is `any`, the types for the arguments will be checked against `any` and the return type will be `any`. Applications of binary and unary operators are treated the same as function applications, with the exception of the `+` operator. For the `+` operator, the following rules are applied, in order of priority:

- If the expression on the left side is of type `number` or `literal number type`, check that the other expression is a success type of `number`. The return type is `number`.
- If the expression on the left side is of type `string` or `literal string type`, check that the other expression is a success type of `string`. The return type is `string`.
- If the expression on the right side is of type `number` or `literal number type`, check that the other expression is a success type of `number`. The return type is `number`.
- If the expression on the right side is of type `string` or `literal string type`, check that the other expression is a success type of `string`. The return type is `string`.
- If the expression on the left side cannot be narrowed to either `number` or `string`, check that both sides are success types of `number | string`. The return type is `number | string`.

$$\frac{\Gamma \vdash E_0 : t_0 \quad \Gamma \vdash E_1 : t_1 \quad t_0 = \textit{number} \vee t_0 = \textit{literal number type} \quad t_1 \wedge \textit{number} \neq \emptyset}{\Gamma \vdash E_0 + E_1 : \textit{number}}$$

$$\frac{\Gamma \vdash E_0 : t_0 \quad \Gamma \vdash E_1 : t_1 \quad t_0 = \textit{string} \vee t_0 = \textit{literal string type} \quad t_1 \wedge \textit{string} \neq \emptyset}{\Gamma \vdash E_0 + E_1 : \textit{string}}$$

$$\begin{array}{c}
\Gamma \vdash E_0 : t_0 \quad \Gamma \vdash E_1 : t_1 \quad t_1 = \text{number} \vee t_1 = \textit{literal number type} \quad t_0 \wedge \text{number} \neq \emptyset \\
\hline
\Gamma \vdash E_0 + E_1 : \text{number} \\
\Gamma \vdash E_0 : t_0 \quad \Gamma \vdash E_1 : t_1 \quad t_1 = \text{string} \vee t_1 = \textit{literal string type} \quad t_0 \wedge \text{string} \neq \emptyset \\
\hline
\Gamma \vdash E_0 + E_1 : \text{string} \\
\Gamma \vdash E_0 : t_0 \quad \Gamma \vdash E_1 : t_1 \quad t_0 \wedge (\text{number} \mid \text{string}) \neq \emptyset \quad t_1 \wedge (\text{number} \mid \text{string}) \neq \emptyset \\
\hline
\Gamma \vdash E_0 + E_1 : \text{number} \mid \text{string}
\end{array}$$

For lambda expressions, we extend  $\Gamma$  with the declared types of all the function parameters, and check the type of the function body against the declared return type. As type syntax is optional, if type annotations are absent for any of the arguments or the return type, the type is assumed to be *any*. The type of the lambda expression is then the function type with the declared types of the parameters and the return type.

$$\begin{array}{c}
\Gamma[x_1 \leftarrow t_1] \cdots [x_n \leftarrow t_n] \Gamma' \quad \Gamma' \vdash S : t' \quad t' \wedge t \neq \emptyset \\
\hline
\Gamma \vdash ((x_1 : t_1, \dots, x_n : t_n) : t \Rightarrow \{ S \}) : (t_1, \dots, t_n) \rightarrow t
\end{array}$$

The type of a conditional expression is the union of the type of its consequent expression and its alternate expression. The predicate expression of a conditional expression must be a success type of a boolean.

$$\begin{array}{c}
\Gamma \vdash E_{pred} : t_{pred} \quad \Gamma \vdash E_{cons} : t_{cons} \quad \Gamma \vdash E_{alt} : t_{alt} \quad t_{pred} \wedge \text{boolean} \neq \emptyset \\
\hline
\Gamma \vdash (E_{pred} ? E_{cons} : E_{alt}) : t_{cons} \mid t_{alt}
\end{array}$$

For *as* expressions, the type to cast the expression to must be a success type of the type of the expression.

$$\begin{array}{c}
\Gamma \vdash E : t' \quad t \wedge t' \neq \emptyset \\
\hline
\Gamma \vdash (E \text{ as } t) : t
\end{array}$$

### 2.3.2 Typing Relations on Statements

For constant declarations, the declared type of  $x$ ,  $t$ , is added to the type environment. As type syntax is optional, if the type annotation for  $x$  is absent, the declared type  $t$  is assumed to be *any*. The derived type of the expression  $E$  must be a success type of the type declared for name  $x$ . The type of the constant declaration statement itself is *undefined*.

$$\begin{array}{c}
\Gamma[x \leftarrow t] \Gamma' \quad \Gamma' \vdash E : t' \quad t' \wedge t \neq \emptyset \\
\hline
\Gamma \vdash (\text{const } x : t = E;) : \text{undefined}
\end{array}$$

For type alias declarations, the declared type of  $x$ ,  $t$ , is added to the type environment. The type of the type alias declaration itself is *undefined*.

$$\begin{array}{c}
\Gamma[x \leftarrow t] \Gamma' \\
\hline
\Gamma \vdash (\text{type } x = t;) : \text{undefined}
\end{array}$$



The type of return statements and expression statements is the type of the expression in the statement.

$$\frac{\Gamma \vdash E : t}{\Gamma \vdash (\text{return } E;) : t} \quad \frac{\Gamma \vdash E : t}{\Gamma \vdash (E;) : t}$$

For blocks,  $\Gamma$  is first extended to include the types of names declared in the block. Then, the component statements are checked against the extended type environment.

We assume that whenever there is a return statement or a conditional statement with a return statement within a block, it is the last statement in the block. (One could consider a “dead code” error otherwise.)

The type of a function body block is the type of the return statement in the block. If the block does not contain any return statements, the type is undefined.

$$\frac{\Gamma[x_1 \leftarrow t_1] \cdots [x_n \leftarrow t_n] \Gamma' \quad \Gamma' \vdash S_1 : t'_1, \dots, \Gamma' \vdash S_n : t'_n \quad S_n \text{ is a return statement}}{\Gamma \vdash \{(\text{const } x_1 : t_1 = E_1;) \dots (\text{const } x_1 : t_1 = E_1;) S_1, \dots, S_n\} : t'_n}$$

$$\frac{\Gamma[x_1 \leftarrow t_1] \cdots [x_n \leftarrow t_n] \Gamma' \quad \Gamma' \vdash S_1 : t'_1, \dots, \Gamma' \vdash S_n : t'_n \quad S_n \text{ is not a return statement}}{\Gamma \vdash \{(\text{const } x_1 : t_1 = E_1;) \dots (\text{const } x_1 : t_1 = E_1;) S_1, \dots, S_n\} : \text{undefined}}$$

The type of a conditional statement or if statement is the union of the type of its consequent statement and its alternate statement. The predicate expression of a conditional statement must be a success type of a boolean.

$$\frac{\Gamma \vdash S_{pred} : t_{pred} \quad \Gamma \vdash S_{cons} : t_{cons} \quad \Gamma \vdash S_{alt} : t_{alt} \quad t_{pred} \wedge \text{boolean} \neq \emptyset}{\Gamma \vdash (\text{if } (S_{pred}) \{ S_{cons} \} \text{ else } \{ S_{alt} \}) : t_{cons} \mid t_{alt}}$$

### 3 Dynamic Type Checking

Expressions evaluate to numbers, boolean values, strings or function values. Implementations of Source generate error messages when unexpected values are used as follows.

Only function values can be applied using the syntax:

$$\text{expression} ::= \text{name}(\text{expressions})$$

For compound functions, implementations need to check that the number of *expressions* matches the number of parameters.

The following table specifies what arguments Source’s operators take and what results they return. Implementations need to check the types of arguments and generate an error message when the types do not match.

operator	argument 1	argument 2	result
+	number	number	number
+	string	string	string
-	number	number	number
*	number	number	number
/	number	number	number
%	number	number	number
===	number	number	bool
===	string	string	bool
!==	number	number	bool
!==	string	string	bool
>	number	number	bool
>	string	string	bool
<	number	number	bool
<	string	string	bool
>=	number	number	bool
>=	string	string	bool
<=	number	number	bool
<=	string	string	bool
&&	bool	any	any
	bool	any	any
!	bool		bool
-	number		number

Preceding `?` and following `if`, Source only allows boolean expressions.

## 4 Standard Libraries

The following libraries are always available in this language.

### MISC Library

The following names are provided by the MISC library:

- `get_time()`: *primitive*, returns number of milliseconds elapsed since January 1, 1970 00:00:00 UTC
- `parse_int(s, i)`: *primitive*, interprets the *string* `s` as an integer, using the positive integer `i` as radix, and returns the respective value, see [ECMAScript Specification, Section 18.2.5](#).
- `undefined`, `NaN`, `Infinity`: *primitive*, refer to JavaScript’s `undefined`, `NaN` (“Not a Number”) and `Infinity` values, respectively.
- `is_boolean(x)`, `is_number(x)`, `is_string(x)`, `is_undefined(x)`, `is_function(x)`: *primitive*, returns `true` if the type of `x` matches the function name and `false` if it does not. Following JavaScript, we specify that `is_number` returns `true` for `NaN` and `Infinity`.
- `prompt(s)`: *primitive*, pops up a window that displays the *string* `s`, provides an input line for the user to enter a text, a “Cancel” button and an “OK” button. The call of `prompt` suspends execution of the program until one of the two buttons is pressed. If the “OK” button is pressed, `prompt` returns the entered text as a string. If the “Cancel” button is pressed, `prompt` returns a non-string value.
- `display(x)`: *primitive*, displays the value `x` in the console<sup>9</sup>; returns the argument `a`.
- `display(x, s)`: *primitive*, displays the string `s`, followed by a space character, followed by the value `x` in the console<sup>9</sup>; returns the argument `x`.
- `error(x)`: *primitive*, displays the value `x` in the console<sup>9</sup> with error flag. The evaluation of any call of `error` aborts the running program immediately.

<sup>9</sup>The notation used for the display of values is consistent with [JSON](#), but also displays `undefined` and function objects.

- `error(x, s)`: *primitive*, displays the string `s`, followed by a space character, followed by the value `x` in the console<sup>9</sup> with error flag. The evaluation of any call of `error` aborts the running program immediately.
- `stringify(x)`: *primitive*, returns a string that represents<sup>9</sup> the value `x`.

All library functions can be assumed to run in  $O(1)$  time, except `display`, `error` and `stringify`, which run in  $O(n)$  time, where  $n$  is the size (number of components such as pairs) of their first argument.

## MATH Library

The following names are provided by the MATH library:

- `math_name`, where *name* is any name specified in the JavaScript Math library, see [ECMAScript Specification, Section 20.2](#). Examples:
  - `math_PI`: *primitive*, refers to the mathematical constant  $\pi$ ,
  - `math_sqrt(n)`: *primitive*, returns the square root of the *number* `n`.

All functions can be assumed to run in  $O(1)$  time and are considered *primitive*.

## List Support

The following list processing functions are supported:

- `pair(x, y)`: *primitive*, makes a pair from `x` and `y`.
- `is_pair(x)`: *primitive*, returns `true` if `x` is a pair and `false` otherwise.
- `head(x)`: *primitive*, returns the head (first component) of the pair `x`.
- `tail(x)`: *primitive*, returns the tail (second component) of the pair `x`.
- `is_null(xs)`: *primitive*, returns `true` if `xs` is the empty list `null`, and `false` otherwise.
- `is_list(x)`: *primitive*, returns `true` if `x` is a list as defined in the lectures, and `false` otherwise. Iterative process; time:  $O(n)$ , space:  $O(1)$ , where  $n$  is the length of the chain of tail operations that can be applied to `x`.
- `list(x1, x2, ..., xn)`: *primitive*, returns a list with  $n$  elements. The first element is `x1`, the second `x2`, etc. Iterative process; time:  $O(n)$ , space:  $O(n)$ , since the constructed list data structure consists of  $n$  pairs, each of which takes up a constant amount of space.
- `draw_data(x1, x2, ..., xn)`: *primitive*, visualizes each `x1, x2, ..., xn` in a separate drawing area in the Source Academy using a box-and-pointer diagram; time, space:  $O(n)$ , where  $n$  is the combined number of data structures such as pairs in `x1, x2, ..., xn`.
- `equal(x1, x2)`: Returns `true` if both have the same structure with respect to `pair`, and the same numbers, boolean values, functions or empty list at corresponding leaf positions (places that are not themselves pairs), and `false` otherwise; time, space:  $O(n)$ , where  $n$  is the number of pairs in `x`.
- `length(xs)`: Returns the length of the list `xs`. Iterative process; time:  $O(n)$ , space:  $O(1)$ , where  $n$  is the length of `xs`.
- `map(f, xs)`: Returns a list that results from list `xs` by element-wise application of `f`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `build_list(f, n)`: Makes a list with  $n$  elements by applying the unary function `f` to the numbers `0` to `n - 1`. Iterative process; time:  $O(n)$ , space:  $O(n)$ .
- `for_each(f, xs)`: Applies `f` to every element of the list `xs`, and then returns `true`. Iterative process; time:  $O(n)$ , space:  $O(1)$ , where  $n$  is the length of `xs`.
- `list_to_string(xs)`: Returns a string that represents list `xs` using the text-based box-and-pointer notation `[...]`.

- `reverse(xs)`: Returns list `xs` in reverse order. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`. The process is iterative, but consumes space  $O(n)$  because of the result list.
- `append(xs, ys)`: Returns a list that results from appending the list `ys` to the list `xs`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `member(x, xs)`: Returns first postfix sublist whose head is identical to `x` (`===`); returns `[]` if the element does not occur in the list. Iterative process; time:  $O(n)$ , space:  $O(1)$ , where  $n$  is the length of `xs`.
- `remove(x, xs)`: Returns a list that results from `xs` by removing the first item from `xs` that is identical (`===`) to `x`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `remove_all(x, xs)`: Returns a list that results from `xs` by removing all items from `xs` that are identical (`===`) to `x`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `filter(pred, xs)`: Returns a list that contains only those elements for which the one-argument function `pred` returns `true`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `enum_list(start, end)`: Returns a list that enumerates numbers starting from `start` using a step size of 1, until the number exceeds (`>`) `end`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`.
- `list_ref(xs, n)`: Returns the element of list `xs` at position `n`, where the first element has index 0. Iterative process; time:  $O(n)$ , space:  $O(1)$ , where  $n$  is the length of `xs`.
- `accumulate(op, initial, xs)`: Applies binary function `op` to the elements of `xs` from right-to-left order, first applying `op` to the last element and the value `initial`, resulting in  $r_1$ , then to the second-last element and  $r_1$ , resulting in  $r_2$ , etc, and finally to the first element and  $r_{n-1}$ , where  $n$  is the length of the list. Thus, `accumulate(op, zero, list(1, 2, 3))` results in `op(1, op(2, op(3, zero)))`. Iterative process; time:  $O(n)$ , space:  $O(n)$ , where  $n$  is the length of `xs`, assuming `op` takes constant time.

## Deviations from JavaScript

We intend the Source language to be a conservative extension of JavaScript: Every correct Source program should behave *exactly* the same using a Source implementation, as it does using a JavaScript implementation. We assume, of course, that suitable libraries are used by the JavaScript implementation, to account for the predefined names of each Source language. This section lists some exceptions where we think a Source implementation should be allowed to deviate from the JavaScript specification, for the sake of internal consistency and esthetics.

**Evaluation result of programs:** JavaScript statically distinguishes between *value-producing* and *non-value-producing statements*. All declarations are non-value-producing, and all expression statements, conditional statements and assignments are value-producing. A block is value-producing if its body statement is value-producing, and then its value is the value of its body statement. A sequence is value-producing if any of its component statements is value-producing, and then its value is the value of its *last* value-producing component statement. The value of an expression statement is the value of the expression. The value of a conditional statement is the value of the branch that gets executed, or the value `undefined` if that branch is not value-producing. The value of an assignment is the value of the expression to the right of its `=` sign. Finally, if the whole program is not value-producing, its value is the value `undefined`.

Example 1:

```
1;
{
  // empty block
}
```

The result of evaluating this program in JavaScript is 1.

Example 2:

```
1;  
{  
  if (true) {} else {}  
}
```

The result of evaluating this program in JavaScript is `undefined`.

Implementations of Source are currently allowed to opt for a simpler scheme.

**Hoisting of function declarations:** In JavaScript, function declarations are “hoisted” (automagically moved) to the beginning of the block in which they appear. This means that applications of functions that are declared with function declaration statements never fail because the name is not yet assigned to their function value. The specification of Source does not include this hoisting; in Source, function declaration can be seen as syntactic sugar for constant declaration and lambda expression. As a consequence, application of functions declared with function declaration may fail in Source if the name that appears as function expression is not yet assigned to the function value it is supposed to refer to.

## Appendix: List library

Those list library functions that are not primitive functions are pre-declared as follows:

```
// list.js START

/**
 * makes a pair whose head (first component) is <CODE>x</CODE>
 * and whose tail (second component) is <CODE>y</CODE>.
 * @param {value} x - given head
 * @param {value} y - given tail
 * @returns {pair} pair with <CODE>x</CODE> as head and <CODE>y</CODE> as tail.
 */
function pair(x, y) {}

/**
 * returns <CODE>>true</CODE> if <CODE>x</CODE> is a
 * pair and false otherwise.
 * @param {value} x - given value
 * @returns {boolean} whether <CODE>x</CODE> is a pair
 */
function is_pair(x) {}

/**
 * returns head (first component) of given pair <CODE>p</CODE>
 * @param {pair} p - given pair
 * @returns {value} head of <CODE>p</CODE>
 */
function head(p) {}

/**
 * returns tail (second component of given pair <CODE>p</CODE>
 * @param {pair} p - given pair
 * @returns {value} tail of <CODE>p</CODE>
 */
function tail(p) {}

/**
 * returns <CODE>>true</CODE> if <CODE>x</CODE> is the
 * empty list <CODE>>null</CODE>, and <CODE>>false</CODE> otherwise.
 * @param {value} x - given value
 * @returns {boolean} whether <CODE>x</CODE> is <CODE>>null</CODE>
 */
function is_null(x) {}

/**
 * Returns <CODE>>true</CODE> if
 * <CODE>xs</CODE> is a list as defined in the textbook, and
 * <CODE>>false</CODE> otherwise. Iterative process;
 * time: <CODE>O(n)</CODE>, space: <CODE>O(1)</CODE>, where <CODE>n</CODE>
 * is the length of the
 * chain of <CODE>tail</CODE> operations that can be applied to <CODE>xs</CODE>.
 * recurses down the list and checks that it ends with the empty list null
 * @param {value} xs - given candidate
 * @returns whether {xs} is a list
 */
function is_list(xs) {}

/**
 * Given <CODE>n</CODE> values, returns a list of length <CODE>n</CODE>.
 * The elements of the list are the given values in the given order.
 */
```

```

* @param {value} value1,value2,...,value_n - given values
* @returns {list} list containing all values
*/
function list(value1, value2, ...values ) {}

/**
* visualizes <CODE>x</CODE> in a separate drawing
* area in the Source Academy using a box-and-pointer diagram; time, space:
* O(n), where n is the total number of data structures such as
* pairs in all the separate structures provided in <CODE>x</CODE>.
* @param {value} value1,value2,...,value_n - given values
* @returns {value} given <CODE>x</CODE>
*/
function draw_data(value1, value2, ...values ) {}

/**
* Returns <CODE>>true</CODE> if both
* have the same structure with respect to <CODE>pair</CODE>,
* and identical values at corresponding leaf positions (places that are not
* themselves pairs), and <CODE>>false</CODE> otherwise. For the "identical",
* the values need to have the same type, otherwise the result is
* <CODE>>false</CODE>. If corresponding leaves are boolean values, these values
* need to be the same. If both are <CODE>undefined</CODE> or both are
* <CODE>>null</CODE>, the result is <CODE>>true</CODE>. Otherwise they are compared
* with <CODE>===</CODE> (using the definition of <CODE>===</CODE> in the
* respective Source language in use). Time, space:
* <CODE>O(n)</CODE>, where <CODE>n</CODE> is the number of pairs in
* <CODE>x</CODE>.
* @param {value} x - given value
* @param {value} y - given value
* @returns {boolean} whether <CODE>x</CODE> is structurally equal to <CODE>y</CODE>
*/
function equal(xs, ys) {
  return is_pair(xs)
    ? (is_pair(ys) &&
       equal(head(xs), head(ys)) &&
       equal(tail(xs), tail(ys)))
    : is_null(xs)
    ? is_null(ys)
    : is_number(xs)
    ? (is_number(ys) && xs === ys)
    : is_boolean(xs)
    ? (is_boolean(ys) && ((xs && ys) || (!xs && !ys)))
    : is_string(xs)
    ? (is_string(ys) && xs === ys)
    : is_undefined(xs)
    ? is_undefined(ys)
    : // we know now that xs is a function
      (is_function(ys) && xs === ys);
}

/**
* Returns the length of the list
* <CODE>xs</CODE>.
* Iterative process; time: <CODE>O(n)</CODE>, space:
* <CODE>O(1)</CODE>, where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
* @param {list} xs - given list
* @returns {number} length of <CODE>xs</CODE>
*/
function length(xs) {

```

```

    return $length(xs, 0);
}
function $length(xs, acc) {
    return is_null(xs) ? acc : $length(tail(xs), acc + 1);
}

/**
 * Returns a list that results from list
 * <CODE>xs</CODE> by element-wise application of unary function <CODE>f</CODE>.
 * Iterative process; time: <CODE>O(n)</CODE>,
 * space: <CODE>O(n)</CODE>, where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * <CODE>f</CODE> is applied element-by-element:
 * <CODE>map(f, list(1, 2))</CODE> results in <CODE>list(f(1), f(2))</CODE>.
 * @param {function} f - unary
 * @param {list} xs - given list
 * @returns {list} result of mapping
 */
function map(f, xs) {
    return $map(f, xs, null);
}
function $map(f, xs, acc) {
    return is_null(xs)
        ? reverse(acc)
        : $map(f, tail(xs), pair(f(head(xs)), acc));
}

/**
 * Makes a list with <CODE>n</CODE>
 * elements by applying the unary function <CODE>f</CODE>
 * to the numbers 0 to <CODE>n - 1</CODE>, assumed to be a nonnegative integer.
 * Iterative process; time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>.
 * @param {function} f - unary function
 * @param {number} n - given nonnegative integer
 * @returns {list} resulting list
 */
function build_list(fun, n) {
    return $build_list(n - 1, fun, null);
}
function $build_list(i, fun, already_built) {
    return i < 0 ? already_built : $build_list(i - 1, fun, pair(fun(i), already_built));
}

/**
 * Applies unary function <CODE>f</CODE> to every
 * element of the list <CODE>xs</CODE>.
 * Iterative process; time: <CODE>O(n)</CODE>, space: <CODE>O(1)</CODE>,
 * Where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * <CODE>f</CODE> is applied element-by-element:
 * <CODE>for_each(fun, list(1, 2))</CODE> results in the calls
 * <CODE>fun(1)</CODE> and <CODE>fun(2)</CODE>.
 * @param {function} f - unary
 * @param {list} xs - given list
 * @returns {boolean} true
 */
function for_each(fun, xs) {
    if (is_null(xs)) {
        return true;
    } else {

```



```

    fun(head(xs));
    return for_each(fun, tail(xs));
  }
}

/**
 * Returns a string that represents
 * list <CODE>xs</CODE> using the text-based box-and-pointer notation
 * <CODE>[...]</CODE>.
 * @param {list} xs - given list
 * @returns {string} <CODE>xs</CODE> converted to string
 */
function list_to_string(xs) {
  return $list_to_string(xs, x => x);
}
function $list_to_string(xs, cont) {
  return is_null(xs)
    ? cont("null")
    : is_pair(xs)
    ? $list_to_string(
        head(xs),
        x => $list_to_string(
            tail(xs),
            y => cont "[" + x + ", " + y + "]"
        ))
    : cont(stringify(xs));
}

/**
 * Returns list <CODE>xs</CODE> in reverse
 * order. Iterative process; time: <CODE>O(n)</CODE>,
 * space: <CODE>O(n)</CODE>, where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * The process is iterative, but consumes space <CODE>O(n)</CODE>
 * because of the result list.
 * @param {list} xs - given list
 * @returns {list} <CODE>xs</CODE> in reverse
 */
function reverse(xs) {
  return $reverse(xs, null);
}
function $reverse(original, reversed) {
  return is_null(original)
    ? reversed
    : $reverse(tail(original), pair(head(original), reversed));
}

/**
 * Returns a list that results from
 * appending the list <CODE>ys</CODE> to the list <CODE>xs</CODE>.
 * Iterative process; time: <CODE>O(n)</CODE>, space:
 * <CODE>O(n)</CODE>, where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * In the result, null at the end of the first argument list
 * is replaced by the second argument, regardless what the second
 * argument consists of.
 * @param {list} xs - given first list
 * @param {list} ys - given second list
 * @returns {list} result of appending <CODE>xs</CODE> and <CODE>ys</CODE>
 */
function append(xs, ys) {
  return $append(xs, ys, xs => xs);
}

```

```

}
function $append(xs, ys, cont) {
  return is_null(xs)
    ? cont(ys)
    : $append(tail(xs), ys, zs => cont(pair(head(xs), zs)));
}

/**
 * Returns first postfix sublist
 * whose head is identical to
 * <CODE>v</CODE> (using <CODE>===</CODE>); returns <CODE>>null</CODE> if the
 * element does not occur in the list.
 * Iterative process; time: <CODE>O(n)</CODE>,
 * space: <CODE>O(1)</CODE>, where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * @param {value} v - given value
 * @param {list} xs - given list
 * @returns {list} postfix sublist that starts with <CODE>v</CODE>
 */
function member(v, xs) {
  return is_null(xs)
    ? null
    : (v === head(xs))
    ? xs
    : member(v, tail(xs));
}

/** Returns a list that results from
 * <CODE>xs</CODE> by removing the first item from <CODE>xs</CODE> that
 * is identical (<CODE>===</CODE>) to <CODE>v</CODE>.
 * Returns the original
 * list if there is no occurrence. Iterative process;
 * time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>, where <CODE>n</CODE>
 * is the length of <CODE>xs</CODE>.
 * @param {value} v - given value
 * @param {list} xs - given list
 * @returns {list} <CODE>xs</CODE> with first occurrence of <CODE>v</CODE> removed
 */
function remove(v, xs) {
  return $remove(v, xs, null);
}
function $remove(v, xs, acc) {
  return is_null(xs)
    ? append(reverse(acc), xs)
    : v === head(xs)
    ? append(reverse(acc), tail(xs))
    : $remove(v, tail(xs), pair(head(xs), acc));
}

/**
 * Returns a list that results from
 * <CODE>xs</CODE> by removing all items from <CODE>xs</CODE> that
 * are identical (<CODE>===</CODE>) to <CODE>v</CODE>.
 * Returns the original
 * list if there is no occurrence.
 * Iterative process;
 * time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>, where <CODE>n</CODE>
 * is the length of <CODE>xs</CODE>.
 * @param {value} v - given value
 * @param {list} xs - given list
 * @returns {list} <CODE>xs</CODE> with all occurrences of <CODE>v</CODE> removed

```

```

*/
function remove_all(v, xs) {
  return $remove_all(v, xs, null);
}
function $remove_all(v, xs, acc) {
  return is_null(xs)
    ? append(reverse(acc), xs)
    : v === head(xs)
    ? $remove_all(v, tail(xs), acc)
    : $remove_all(v, tail(xs), pair(head(xs), acc));
}

/**
 * Returns a list that contains
 * only those elements for which the one-argument function
 * <CODE>pred</CODE>
 * returns <CODE>>true</CODE>.
 * Iterative process;
 * time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>,
 * where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
 * @param {function} pred - unary function returning boolean value
 * @param {list} xs - given list
 * @returns {list} list with those elements of <CODE>xs</CODE> for which <CODE>pred</CODE>
 */
function filter(pred, xs) {
  return $filter(pred, xs, null);
}
function $filter(pred, xs, acc) {
  return is_null(xs)
    ? reverse(acc)
    : pred(head(xs))
    ? $filter(pred, tail(xs), pair(head(xs), acc))
    : $filter(pred, tail(xs), acc);
}

/**
 * Returns a list that enumerates
 * numbers starting from <CODE>start</CODE> using a step size of 1, until
 * the number exceeds (<CODE>></CODE>) <CODE>end</CODE>.
 * Iterative process;
 * time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>,
 * where <CODE>n</CODE> is <CODE>end - start</CODE>.
 * @param {number} start - starting number
 * @param {number} end - ending number
 * @returns {list} list from <CODE>start</CODE> to <CODE>end</CODE>
 */
function enum_list(start, end) {
  return $enum_list(start, end, null);
}
function $enum_list(start, end, acc) {
  return start > end
    ? reverse(acc)
    : $enum_list(start + 1, end, pair(start, acc));
}

/**
 * Returns the element
 * of list <CODE>xs</CODE> at position <CODE>n</CODE>,
 * where the first element has index 0.
 * Iterative process;

```

```

* time: <CODE>O(n)</CODE>, space: <CODE>O(1)</CODE>,
* where <CODE>n</CODE> is the length of <CODE>xs</CODE>.
* @param {list} xs - given list
* @param {number} n - given position
* @returns {value} item in <CODE>xs</CODE> at position <CODE>n</CODE>
*/
function list_ref(xs, n) {
  return n === 0
    ? head(xs)
    : list_ref(tail(xs), n - 1);
}

/** Applies binary
* function <CODE>f</CODE> to the elements of <CODE>xs</CODE> from
* right-to-left order, first applying <CODE>f</CODE> to the last element
* and the value <CODE>initial</CODE>, resulting in <CODE>r</CODE><SUB>1</SUB>,
* then to the
* second-last element and <CODE>r</CODE><SUB>1</SUB>, resulting in
* <CODE>r</CODE><SUB>2</SUB>,
* etc, and finally
* to the first element and <CODE>r</CODE><SUB>n-1</SUB>, where
* <CODE>n</CODE> is the length of the
* list. Thus, <CODE>accumulate(f, zero, list(1, 2, 3))</CODE> results in
* <CODE>f(1, f(2, f(3, zero)))</CODE>.
* Iterative process;
* time: <CODE>O(n)</CODE>, space: <CODE>O(n)</CODE>,
* where <CODE>n</CODE> is the length of <CODE>xs</CODE>
* assuming <CODE>f</CODE> takes constant time.
* @param {function} f - binary function
* @param {value} initial - initial value
* @param {list} xs - given list
* @returns {value} result of accumulating <CODE>xs</CODE> using <CODE>f</CODE> starting
*/
function accumulate(f, initial, xs) {
  return $accumulate(f, initial, xs, x => x);
}
function $accumulate(f, initial, xs, cont) {
  return is_null(xs)
    ? cont(initial)
    : $accumulate(f, initial, tail(xs), x => cont(f(head(xs), x)));
}

/**
* Optional second argument.
* Similar to <CODE>display</CODE>, but formats well-formed lists nicely if detected.
* @param {value} xs - list structure to be displayed
* @param {string} s to be displayed, preceding <CODE>xs</CODE>
* @returns {value} xs, the first argument value
*/
function display_list(xs, s) {}

//
// list.js END

```