Cg Language Specification

This is version 2.0 of the Cg Language specification. This language specification describes version 2.0 of the Cg language

Language Overview

The Cg language is primarily modeled on ANSI C, but adopts some ideas from modern languages such as C++ and Java, and from earlier shading languages such as RenderMan and the Stanford shading language. The language also introduces a few new ideas. In particular, it includes features designed to represent data flow in stream-processing architectures such as GPUs. Profiles, which are specified at compile time, may subset certain features of the language, including the ability to implement loops and the precision at which certain computations are performed.

Like C, Cg is designed primarily as a low-level programming language. Features are provided that map as directly as possible to hardware capabilities. Higher level abstractions are designed primarily to not get in the way of writing code that maps directly to the hardware in the most efficient way possible. The changes in the language from C primarily reflect differences in the way GPU hardware works compared to conventional CPUs. GPUs are designed to run large numbers of small threads of processing in parallel, each running a copy of the same program on a different data set.

Differences from ANSI C

Cg was developed based on the ANSI-C language with the following major additions, deletions, and changes. (This is a summary-more detail is provided later in this document):

Silent Incompatibilities

Most of the changes from ANSI C are either omissions or additions, but there are a few potentially silent incompatibilities. These are changes within Cg that could cause a program that compiles without errors to behave in a manner different from C:

- The type promotion rules for constants are different when the constant is not explicitly typed using a type cast or type suffix. In general, a binary operation between a constant that is not explicitly typed and a variable is performed at the variable's precision, rather than at the constant's default precision.
- Declarations of `struct` perform an automatic `typedef` (as in C++) and thus could override a previously declared type.
- Arrays are first-class types that are distinct from pointers. As a result, array assignments semantically perform a copy operation for the entire array.

Similar Operations That Must be Expressed Differently

There are several changes that force the same operation to be expressed differently in Cg than in C:

- A Boolean type, `bool`, is introduced, with corresponding implications for operators and control constructs.
- Arrays are first-class types because Cg does not support pointers.
- Functions pass values by value/result, and thus use an `out` or `inout` modifier in the formal parameter list to return a parameter. By default, formal parameters are `in`, but it is acceptable to specify this explicitly. Parameters can also be specified as `in out`, which is semantically the same as `inout`.

C features not present in Cg

- Language profiles (described in the Profiles section) may subset language capabilities in a variety of ways. In particular, language profiles may restrict the use of for and while loops. For example, some profiles may only support loops that can be fully unrolled at compile time.

- Reserved keywords `goto`, `switch`, `case`, and `default` are not supported, nor are labels.
- Pointers and pointer-related capabilities, such as the `&` and `->` operators, are not supported.
- Arrays are supported, but with some limitations on size and dimensionality. Restrictions on the use of computed subscripts are also permitted. Arrays may be designated as `packed`. The operations allowed on packed arrays may be different from those allowed on unpacked arrays. Predefined `packed` types are provided for vectors and matrices. It is strongly recommended that these predefined types be used.
- There is no `enum` or `union`.
- There are no bit-field declarations in structures.
- All integral types are implicitly signed, there is no *signed* keyword.

Cg features not present in C

- A *binding semantic* may be associated with a structure tag, a variable, or a structure element to denote that object's mapping to a specific hardware or API resource. Binding semantics are described in the *Binding Semantics* section.
- There is a built-in swizzle operator: `.xyzw` or `.rgba` for vectors. This operator allows the components of a vector to be rearranged and also replicated. It also allows the creation of a vector from a scalar.
- For an lvalue, the swizzle operator allows components of a vector or matrix to be selectively written.
- There is a similar built-in swizzle operator for matrices: `._m<row><col>[_m<row><col>][...]`. This operator allows access to individual matrix components and allows the creation of a vector from elements of a matrix. For compatibility with DirectX 8 notation, there is a second form of matrix swizzle, which is described later.
- Numeric data types are different. Cg's primary numeric data types are `float`, `half`, and `fixed`. Fragment profiles are required to support all three data types, but may choose to implement `half` and/or `fixed` at `float` precision. Vertex profiles are required to support `half` and `float`, but may choose to implement `half` at `float` precision. Vertex profiles may omit support for `fixed` operations, but must still support definition of `fixed` variables. Cg allows profiles to omit run-time support for `int` and other integer types. Cg allows profiles to treat `double` as `float`.
- Many operators support per-element vector operations.
- The `?:`, `||`, `&&`, `!`, and comparison operators can be used with `bool` vectors to perform multiple conditional operations simultaneously.

  The side effects of all operands to vector `?:`, `||`, and `&&` operators are always executed.
- Non-static global variables, and parameters to top-level functions (such as *main()*) may be designated as `uniform`. A `uniform` variable may be read and written within a program, just like any other variable. However, the uniform modifier indicates that the initial value of the variable/parameter is expected to be constant across a large number of invocations of the program.
- A new set of `sampler*` types represents handles to texture sampler units.
- Functions may have default values for their parameters, as in C++. These defaults are expressed using assignment syntax.
- Function and operator overloading is supported.
- Variables may be defined anywhere before they are used, rather than just at the beginning of a scope as in C. (That is, we adopt the C++ rules that govern where variable declarations are allowed.) Variables may not be redeclared within the same scope.

- Vector constructors, such as the form `float4(1,2,3,4)`, and matrix constructors may be used anywhere in an expression.
- A `struct` definition automatically performs a corresponding `typedef`, as in C++.
- C++–style `//` comments are allowed in addition to C-style `/*` ... `*/` comments.
- A limited form of inheritance is supported; `interface` types may be defined which contain only member functions (no data members) and `struct` types may inherit from a single interface and provide specific implementations for all the member functions. Interface objects may not be created; a variable of interface type may have any implementing struct type assigned to it.

Detailed Language Specification

Definitions

The following definitions are based on the ANSI C standard:

Object:
: An object is a region of data storage in the execution environment, the contents of which can represent values. When referenced, an object may be interpreted as having a particular type.

Declaration:
: A declaration specifies the interpretation and attributes of a set of identifiers.

Definition:
: A declaration that also causes storage to be reserved for an object or code that will be generated for a function named by an identifier is a definition.

Profiles

Compilation of a Cg program, a top-level function, always occurs in the context of a compilation profile. The profile specifies whether certain optional language features are supported. These optional language features include certain control constructs and standard library functions. The compilation profile also defines the precision of the `float`, `half`, and `fixed` data types, and specifies whether the `fixed` and `sampler*` data types are fully or only partially supported. The profile also specifies the environment in which the program will be run. The choice of a compilation profile is made externally to the language, by using a compiler command-line switch, for example.

The profile restrictions are only applied to the top-level function that is being compiled and to any variables or functions that it references, either directly or indirectly. If a function is present in the source code, but not called directly or indirectly by the top-level function, it is free to use capabilities that are not supported by the current profile.

The intent of these rules is to allow a single Cg source file to contain many different top-level functions that are targeted at different profiles. The core Cg language specification is sufficiently complete to allow all of these functions to be parsed. The restrictions provided by a compilation profile are only needed for code generation, and are therefore only applied to those functions for which code is being generated. This specification uses the word ‘‘program’’ to refer to the top-level function, any functions the top-level function calls, and any global variables or typedef definitions it references.

Each profile must have a separate specification that describes its characteristics and limitations.

This core Cg specification requires certain minimum capabilities for all profiles. In some cases, the core specification distinguishes between vertex-program and fragment-program profiles, with different minimum capabilities for each.

Declarations and declaration specifiers.

A Cg program consists of a series of declarations, each of which declares one or more variables or functions, or declares and defines a single function. Each declaration consists of zero or more declaration specifiers, a type, and one or more declarators. Some of the declaration specifiers are the same as those in ANSI C; others are new to Cg

const
Marks a variable as a constant that cannot be assigned to within the program. Unless this is combined with `uniform` or `varying`, the declarator must include an initializer to give the variable a value.

extern
Marks this declaration as solely a declaration and not a definition. There must be a non-`extern` declaration elsewhere in the program.

in Only usable on parameter and `varying` declarations. Marks the parameter or varying as an input to the function or program. Function parameters with no `in`, `out`, or `inout` specifier are implicitly `in`

inline
Only usable on a function definition. Tells the compiler that it should always inline calls to the function if at all possible.

inout
Only usable on parameter and `varying` declarations. Marks the parameter or varying as both an input to and an output from the function or program

static
Only usable on global variables. Marks the variable as 'private' to the program, and not visible externally. Cannot be combined with `uniform` or `varying`

out Only usable on parameter and `varying` declarations. Marks the parameter or varying as an output from the function or program

uniform
Only usable on global variables and parameters to the top-level main function of a program. If specified on a non-top-level function parameter it is ignored. The intent of this rule is to allow a function to serve as either a top-level function or as one that is not.

Note that `uniform` variables may be read and written just like non-`uniform` variables. The `uniform` qualifier simply provides information about how the initial value of the variable is to be specified and stored, through a mechanism external to the language.

varying
Only usable on global variables and parameters to the top-level main function of a program. If specified on a non-top-level function parameter it is ignored.

profile name
The name of any profile (or profile wildcard — see Profiles) may be used as a specifier on any function declaration. It defines a function that is only visible in the corresponding profiles.

The specifiers `uniform` and `varying` specify how data is transferred between the rest of the world and a Cg program. Typically, the initial value of a `uniform` variable or parameter is stored in a different class of hardware register for a `varying`. Furthermore, the external mechanism for specifying the initial value of `uniform` variables or parameters may be different than that used for specifying the initial value of `varying` variables or parameters. Parameters qualified as `uniform` are normally treated as persistent state, while `varying` parameters are treated as streaming data, with a new value specified for each stream record (such as within a vertex array).

Non-`static` global variables are treated as `uniform` by default, while parameters to the top-level function are treated as `varying` by default.

Each declaration is visible (''in scope'') from the point of its declarator until the end of the enclosing block or the end of the compilation unit if outside any block. Declarations in named scopes (such as structs and interfaces) may be visible outside of their scope using explicit scope qualifiers, as in C++.

Semantics

Each declarator in a declaration may optionally have a semantic specified with it. A semantic specifies how the variable is connected to the environment in which the program runs. All semantics are profile specific (so they have different meanings in different profiles), though there is some attempt to be consistent across profiles. Each profile specification must specify the set of semantics which the profile understands, as well as what behavior occurs for any other unspecified semantics.

Function Declarations

Functions are declared essentially as in C. A function that does not return a value must be declared with a `void` return type. A function that takes no parameters may be declared in one of two ways:

As in C, using the void keyword:

```
functionName(void)
```

With no parameters at all:

```
functionName()
```

Functions may be declared as `static`. If so, they may not be compiled as a program and are not visible externally

Function overloading and optional arguments

Cg supports function overloading; that is you may define multiple functions with the same name. The function actually called at any given call site is based on the types of the arguments at that call site; the definition that best matches is called. See the the Overload resolution entry elsewhere in this document section for the precise rules. Trailing arguments with initializers are optional arguments; defining a function with optional arguments is equivalent to defining multiple overloaded functions that differ by having and not having the optional argument. The value of the initializer is used only for the version that does not have the argument and is ignored if the argument is present.

Overloading of Functions by Profile

Cg supports overloading of functions by compilation profile. This capability allows a function to be implemented differently for different profiles. It is also useful because different profiles may support different subsets of the language capabilities, and because the most efficient implementation of a function may be different for different profiles.

The profile name must precede the return type name in the function declaration. For example, to define two different versions of the function `myfunc` for the `profileA` and `profileB` profiles:

```
profileA float myfunc(float x) {...};
profileB float myfunc(float x) {...};
```

If a type is defined (using a `typedef`) that has the same name as a profile, the identifier is treated as a type name, and is not available for profile overloading at any subsequent point in the file.

If a function definition does not include a profile, the function is referred to as an ‘‘open-profile’’ function. Open-profile functions apply to all profiles.

Several wildcard profile names are defined. The name `vs` matches any vertex profile, while the name `ps` matches any fragment or pixel profile. The names `ps_1` and `ps_2` match any DX8 pixel shader 1.x profile, or DX9 pixel shader 2.x profile, respectively. Similarly, the names `vs_1` and `vs_2` match any DX vertex shader 1.x or 2.x, respectively. Additional valid wildcard profile names may be defined by individual profiles.

In general, the most specific version of a function is used. More details are provided in the section on function overloading, but roughly speaking, the search order is the following:

1. version of the function with the exact profile overload
2. version of the function with the most specific wildcard profile overload (e.g. `vs`, ‘‘ps_1’’)
3. version of function with no profile overload

This search process allows generic versions of a function to be defined that can be overridden as needed for particular hardware.

Syntax for Parameters in Function Definitions

Functions are declared in a manner similar to C, but the parameters in function definitions may include a binding semantic (discussed later) and a default value.

Each parameter in a function definition takes the following form:

```
<declspecs> <type> identifier [: <binding_semantic>] [= <default>]
```

<default> is an expression that resolves to a constant at compile time.

Default values are only permitted for `uniform` parameters, and for `in` parameters to non top-level functions.

Function Calls

A function call returns an rvalue. Therefore, if a function returns an array, the array may be read but not written. For example, the following is allowed:

```
y = myfunc(x)[2];
```

But, this is not:

```
myfunc(x)[2] = y;
```

For multiple function calls within an expression, the calls can occur in any order — it is undefined.

Types

Cg’s types are as follows:

- The `int` type is preferably 32−bit two’s complement. Profiles may optionally treat `int` as `float`.
- The `unsigned` type is preferably a 32−bit ordinal value. `unsigned` may also be used with other integer types to make different sized unsigned values
- The `char`, `short`, and `long` types are two’s complement integers of various sizes. The only requirement is that `char` is no larger that `short`, `short` is no larger than `int` and `long` is at least as large as `int`
- The `float` type is as close as possible to the IEEE single precision (32−bit) floating point format. Profiles must support the `float` data type.
- The `half` type is lower-precision IEEE-like floating point. Profiles must support the `half` type, but may choose to implement it with the same precision as the `float` type.
- The `fixed` type is a signed type with a range of at least [−2,2) and with at least 10 bits of fractional precision. Overflow operations on the data type clamp rather than wrap. Fragment profiles must support the `fixed` type, but may implement it with the same precision as the `half` or `float` types. Vertex profiles are required to provide partial support (as defined below) for the `fixed` type. Vertex profiles have the option to provide full support for the `fixed` type or to implement the `fixed` type with the same precision as the `half` or `float` types.
- The `bool` type represents Boolean values. Objects of `bool` type are either true or false.

- The `cint` type is 32−bit two's complement. This type is meaningful only at compile time; it is not possible to declare objects of type `cint`.
- The `cfloat` type is IEEE single-precision (32−bit) floating point. This type is meaningful only at compile time; it is not possible to declare objects of type `cfloat`.
- The `void` type may not be used in any expression. It may only be used as the return type of functions that do not return a value.
- The `sampler*` types are handles to texture objects. Formal parameters of a program or function may be of type `sampler*`. No other definition of `sampler*` variables is permitted. A `sampler*` variable may only be used by passing it to another function as an `in` parameter. Assignment to `sampler*` variables is not permitted, and `sampler*` expressions are not permitted.

  The following sampler types are always defined: `sampler`, `sampler1D`, `sampler2D`, `sampler3D`, `samplerCUBE`, `samplerRECT`.

  The base `sampler` type may be used in any context in which a more specific sampler type is valid. However, a `sampler` variable must be used in a consistent way throughout the program. For example, it cannot be used in place of both a `sampler1D` and a `sampler2D` in the same program. The `sampler` type is deprecated and only provided for backwards compatibility with Cg 1.0

  Fragment profiles are required to fully support the `sampler`, `sampler1D`, `sampler2D`, `sampler3D`, and `samplerCUBE` data types. Fragment profiles are required to provide partial support (as defined below) for the `samplerRECT` data type and may optionally provide full support for this data type.

  Vertex profiles are required to provide partial support for the six sampler data types and may optionally provide full support for these data types.
- An *array* type is a collection of one or more elements of the same type. An *array* variable has a single index.
- Some array types may be optionally designated as `packed`, using the `packed` type modifier. The storage format of a `packed` type may be different from the storage format of the corresponding unpacked type. The storage format of packed types is implementation dependent, but must be consistent for any particular combination of compiler and profile. The operations supported on a packed type in a particular profile may be different than the operations supported on the corresponding unpacked type in that same profile. Profiles may define a maximum allowable size for packed arrays, but must support at least size 4 for packed vector (1D array) types, and 4x4 for packed matrix (2D array) types.
- When declaring an array of arrays in a single declaration, the `packed` modifier refers to all of the arrays. However, it is possible to declare an unpacked array of `packed` arrays by declaring the first level of array in a `typedef` using the `packed` keyword and then declaring an array of this type in a second statement. It is not possible to declare a packed array of unpacked arrays.
- For any supported numeric data type *TYPE*, implementations must support the following packed array types, which are called *vector types*. Type identifiers must be predefined for these types in the global scope:

```
typedef packed TYPE TYPE1[1];
typedef packed TYPE TYPE2[2];
typedef packed TYPE TYPE3[3];
typedef packed TYPE TYPE4[4];
```

  For example, implementations must predefine the type identifiers `float1`, `float2`, `float3`, `float4`, and so on for any other supported numeric type.

- For any supported numeric data type *TYPE*, implementations must support the following packed array types, which are called *matrix types*. Implementations must also predefine type identifiers (in the global scope) to represent these types:

```
packed TYPE1 TYPE1x1[1];
packed TYPE2 TYPE1x2[1];
packed TYPE3 TYPE1x3[1];
packed TYPE4 TYPE1x4[1];
packed TYPE1 TYPE2x1[2];
packed TYPE2 TYPE2x2[2];
packed TYPE3 TYPE2x3[2];
packed TYPE4 TYPE2x4[2];
packed TYPE1 TYPE3x1[3];
packed TYPE2 TYPE3x2[3];
packed TYPE3 TYPE3x3[3];
packed TYPE4 TYPE3x4[3];
packed TYPE1 TYPE4x1[4];
packed TYPE2 TYPE4x2[4];
packed TYPE3 TYPE4x3[4];
packed TYPE4 TYPE4x4[4];
```

  For example, implementations must predefine the type identifiers `float2x1`, `float3x3`, `float4x4`, and so on. A typedef follows the usual matrix-naming convention of `TYPEr-ows_X_columns`. If we declare `float4x4 a`, then

```
a[3] is equivalent to a._m30_m31_m32_m33
```

  Both expressions extract the third row of the matrix.

- Implementations are required to support indexing of vectors and matrices with constant indices.
- A `struct` type is a collection of one or more members of possibly different types. It may include both function members (methods) and data members (fields).

Struct and Interface types

Interface types are defined with a *interface* keyword in place of the normal *struct* keyword. Interface types may only declare member functions, not data members. Interface member functions may only be declared, not defined (no default implementations in C++ parlance).

Struct types may inherit from a single interface type, and must define an implementation member function for every member function declared in the interface type.

Partial Support of Types

This specification mandates ‘‘partial support’’ for some types. Partial support for a type requires the following:

- Definitions and declarations using the type are supported.
- Assignment and copy of objects of that type are supported (including implicit copies when passing function parameters).
- Top-level function parameters may be defined using that type.

If a type is partially supported, variables may be defined using that type but no useful operations can be performed on them. Partial support for types makes it easier to share data structures in code that is targeted at different profiles.

Type Categories

- The *signed integral* type category includes types `cint`, `char`, `short`, `int`, and `long`.
- The *unsigned integral* type category includes types `unsigned char`, `unsigned short`, `unsigned int`, and `unsigned long`. `unsigned` is the same as `unsigned int`
- The *integral* category includes both *signed integral* and *unsigned integral* types
- The *floating* type category includes types `cfloat`, `float`, `half`, and `fixed` (Note that floating really means floating or fixed/fractional.)
- The *numeric* type category includes *integral* and *floating* types.
- The *compile-time* type category includes types `cfloat` and `cint`. These types are used by the compiler for constant type conversions.
- The *dynamic* type category includes all interface and the unsized array entry elsewhere in this document types
- The *concrete* type category includes all types that are not included in the *compile-time* and *dynamic* type category.
- The *scalar* type category includes all types in the numeric category, the `bool` type, and all types in the compile-time category. In this specification, a reference to a <category> type (such as a reference to a numeric type) means one of the types included in the category (such as `float`, `half`, or `fixed`).

Constants

Constant literals are defined as in C, including an optional `0` or `0x` prefix for octal or hexadecimal constants, and `e` exponent suffix for floating point constants. A constant may be explicitly typed or implicitly typed. Explicit typing of a constant is performed, as in C, by suffixing the constant with a one or two characters indicating the type of the constant:

- **d** for `double`
- **f** for `float`
- **h** for `half`
- **i** for `int`
- **l** for `long`
- **s** for `short`
- **t** for `char`
- **u** for `unsigned`, which may also be followed by **s**, **t**, **i**, or **l**
- **x** for `fixed`

Any constant that is not explicitly typed is implicitly typed. If the constant includes a decimal point or an 'e' exponent suffix, it is implicitly typed as `cfloat`. If it does not include a decimal point, it is implicitly typed as `cint`.

By default, constants are base 10. For compatibility with C, integer hexadecimal constants may be specified by prefixing the constant with `0x`, and integer octal constants may be specified by prefixing the constant with `0`.

Compile-time constant folding is preferably performed at the same precision that would be used if the operation were performed at run time. Some compilation profiles may allow some precision flexibility for the hardware; in such cases the compiler should ideally perform the constant folding at the highest hardware precision allowed for that data type in that profile.

If constant folding cannot be performed at run-time precision, it may optionally be performed using the precision indicated below for each of the numeric datatypes:

float
: s23e8 (‘‘fp32’’) IEEE single precision floating point

half
: s10e5 (‘‘fp16’’) floating point w/ IEEE semantics

fixed
: S1.10 fixed point, clamping to [−2, 2)

double
: s52e11 (‘‘fp64’’) IEEE double precision floating point

int
: signed 32 bit twos-complement integer

char
: signed 8 bit twos-complement integer

short
: signed 16 bit twos-complement integer

long
: signed 64 bit twos-complement integer

Type Conversions

Some type conversions are allowed implicitly, while others require an cast. Some implicit conversions may cause a warning, which can be suppressed by using an explicit cast. Explicit casts are indicated using C-style syntax (e.g., casting `variable` to the `float4` type may be achieved via ‘‘(float4)variablename’’).

Scalar conversions:
: Implicit conversion of any scalar numeric type to any other scalar numeric type is allowed. A warning may be issued if the conversion is implicit and it is possible that precision is lost. implicit conversion of any scalar object type to any compatible scalar object type is also allowed. Conversions between incompatible scalar object types or object and numeric types are not allowed, even with an explicit cast. ‘‘sampler’’ is compatible with ‘‘sampler1D’’, ‘‘sampler2D’’, ‘‘sampler3D’’, ‘‘samplerCube’’, and ‘‘samplerRECT’’. No other object types are compatible (‘‘sampler1D’’ is not compatible with ‘‘sampler2D’’, even though both are compatible with ‘‘sampler’’).

  Scalar types may be implicitly converted to vectors and matrixes of compatible type. The scalar will be replicated to all elements of the vector or matrix. Scalar types may also be explicitly cast to structure types if the scalar type can be legally cast to every member of the structure.

Vector conversions
: Vectors may be converted to scalar types (selects the first element of the vector). A warning is issued if this is done implicitly. A vector may also be implicitly converted to another vector of the same size and compatible element type.

  A vector may be converted to a smaller compatible vector, or a matrix of the same total size, but a warning if issued if an explicit cast is not used.

Matrix conversions
: Matrixes may be converted to a scalar type (selects to 0,0 element). As with vectors, this causes a warning if its done implicitly. A matrix may also be converted implicitly to a matrix of the same size and shape and compatible element type

  A Matrix may be converted to a smaller matrix type (selects the upper- left submatrix), or to a vector of the same total size, but a warning is issued if an explicit cast is not used.

Structure conversions
: a structure may be explicitly cast to the type of its first member, or to another structure type with the same number of members, if each member of the struct can be converted to the corresponding member of the new struct. No implicit conversions of struct types are allowed.

Array conversions

An array may be explicitly converted to another array type with the same number of elements and a compatible element type. A compatible element type is any type to which the element type of the initial array may be implicitly converted to. No implicit conversions of array types are allowed.

Target type	Source: Scalar	Vector	Matrix	Struct	Array
Scalar	A	W	W	E(3)	-
Vector	A	A/W(1)	W(2)	E(3)	E(6)
Matrix	A	W(2)	A/W(1)	E(3)	E(7)
Struct	E	E(4)	E(4)	E(4/5)	E(4)
Array	-	E(6)	E(7)	E(3)	E(6)

A = allowed implicitly or explicitly
W = allowed, but warning issued if implicit
E = only allowed with explicit cast
- = not allowed

notes

(1) not allowed if target is larger than source. Warning if target is smaller than source

(2) only allowed if source and target are the same total size

(3) only if the first member of the source can be converted to the target

(4) only if the target struct contains a single field of the source type

(5) only if both source and target have the same number of members and each member of the source can be converted to the corresponding member of the target.

(6) Source and target sizes must be the same and element types must be compatible

(7) Array type must be an array of vectors that matches the matrix type.

Explicit casts are:

- compile-time type when applied to expressions of compile-time type.
- numeric type when applied to expressions of numeric or compile-time types.
- numeric vector type when applied to another vector type of the same number of elements.
- numeric matrix type when applied to another matrix type of the same number of rows and columns.

Type Equivalency

Type T1 is equivalent to type T2 if any of the following are true:

- T2 is equivalent to T1.

- T1 and T2 are the same scalar, vector, or structure type.

  A packed array type is *not* equivalent to the same size unpacked array.
- T1 is a typedef name of T2.
- T1 and T2 are arrays of equivalent types with the same number of elements.
- The unqualified types of T1 and T2 are equivalent, and both types have the same qualifications.
- T1 and T2 are functions with equivalent return types, the same number of parameters, and all corresponding parameters are pair-wise equivalent.

Type-Promotion Rules

The `cfloat` and `cint` types behave like `float` and `int` types, except for the usual arithmetic conversion behavior (defined below) and function-overloading rules (defined later).

The *usual arithmetic conversions* for binary operators are defined as follows:

1. If one operand is `cint` it is converted to the other type
2. If one operand is `cfloat` and the other is *floating*, the `cfloat` is converted to the other type
3. If both operands are *floating* then the smaller type is converted to the larger type
4. If one operand is *floating* and the other is *integral*, the integral argument is converted to the floating type.
5. If both operands are *integral* the smaller type is converted to the larger type
6. If one operand is *signed integral* while the other is *unsigned integral* and they are the same size, the signed type is converted to unsigned.

Note that conversions happen prior to performing the operation.

Assignment

Assignment of an expression to a concrete typed object converts the expression to the type of the object. The resulting value is then assigned to the object or value.

The value of the assignment expressions (`=`, `*=`, and so on) is defined as in C:

An assignment expression has the value of the left operand after the assignment but is not an lvalue. The type of an assignment expression is the type of the left operand unless the left operand has a qualified type, in which case it is the unqualified version of the type of the left operand. The side effect of updating the stored value of the left operand occurs between the previous and the next sequence point.

An assignment of an expression to a dynamic typed object is only possible if the type of the expression is compatible with the dynamic object type. The object will then take on the type of the expression assigned to it until the next assignment to it.

“Smearing” of Scalars to Vectors

If a binary operator is applied to a vector and a scalar, the scalar is automatically type-promoted to a same-sized vector by replicating the scalar into each component. The ternary `?:` operator also supports smearing. The binary rule is applied to the second and third operands first, and then the binary rule is applied to this result and the first operand.

Namespaces

Just as in C, there are two namespaces. Each has multiple scopes, as in C.

- Tag namespace, which consists of `struct` tags

- Regular namespace:
  - typedef names (including an automatic `typedef` from a `struct` declaration)
  - variables
  - function names

Arrays and Subscripting

Arrays are declared as in C, except that they may optionally be declared to be `packed`, as described earlier. Arrays in Cg are first-class types, so array parameters to functions and programs must be declared using array syntax, rather than pointer syntax. Likewise, assignment of an *array*−typed object implies an array copy rather than a pointer copy.

Arrays with size `[1]` may be declared but are considered a different type from the corresponding non-array type.

Because the language does not currently support pointers, the storage order of arrays is only visible when an application passes parameters to a vertex or fragment program. Therefore, the compiler is currently free to allocate temporary variables as it sees fit.

The declaration and use of arrays of arrays is in the same style as in C. That is, if the 2D array `A` is declared as

```
float A[4][4];
```

then, the following statements are true:

- The array is indexed as `A[row][column];`
- The array can be built with a constructor using

```
float4x4 A = { { A[0][0], A[0][1], A[0][2], A[0][3] },
               { A[1][0], A[1][1], A[1][2], A[1][3] },
               { A[2][0], A[2][1], A[2][2], A[2][3] },
               { A[3][0], A[3][1], A[3][2], A[3][3] } };
```

- `A[0]` is equivalent to `float4(A[0][0], A[0][1], A[0][2], A[0][3])`

Support must be provided for structs containing arrays.

Unsized Arrays

Objects may be declared as *unsized* arrays by using a declaration with an empty size `[]` and no initializer. If a declarator uses unsized array syntax with an initializer, it is declared with a concrete (sized) array type based on the declarator. Unsized arrays are dynamic typed objects that take on the size of any array assigned to them.

Minimum Array Requirements

Profiles are required to provide partial support for certain kinds of arrays. This partial support is designed to support vectors and matrices in all profiles. For vertex profiles, it is additionally designed to support arrays of light state (indexed by light number) passed as uniform parameters, and arrays of skinning matrices passed as uniform parameters.

Profiles must support subscripting, copying, size querying and swizzling of vectors and matrices. However, subscripting with run-time computed indices is not required to be supported.

Vertex profiles must support the following operations for any non-packed array that is a uniform parameter to the program, or is an element of a structure that is a uniform parameter to the program. This requirement also applies when the array is indirectly a uniform program parameter (that is, it and or the structure containing it has been passed via a chain of `in` function parameters). The three operations that must be supported are

- rvalue subscripting by a run-time computed value or a compile-time value.
- passing the entire array as a parameter to a function, where the corresponding formal function parameter is declared as `in`.
- querying the size of the array with a `.length` suffix.

The following operations are explicitly not required to be supported:

- lvalue-subscripting
- copying
- other operators, including multiply, add, compare, and so on

Note that when a uniform array is rvalue subscripted, the result is an expression, and this expression is no longer considered to be a `uniform` program parameter. Therefore, if this expression is an array, its subsequent use must conform to the standard rules for array usage.

These rules are not limited to arrays of numeric types, and thus imply support for arrays of struct, arrays of matrices, and arrays of vectors when the array is a `uniform` program parameter. Maximum array sizes may be limited by the number of available registers or other resource limits, and compilers are permitted to issue error messages in these cases. However, profiles must support sizes of at least `float arr[8]`, `float4 arr[8]`, and `float4x4 arr[4][4]`.

Fragment profiles are not required to support any operations on arbitrarily sized arrays; only support for vectors and matrices is required.

Function Overloading

Multiple functions may be defined with the same name, as long as the definitions can be distinguished by unqualified parameter types and do not have an open-profile conflict (as described in the section on open functions).

Function-matching rules:

1. Add all visible functions with a matching name in the calling scope to the set of function candidates.
2. Eliminate functions whose profile conflicts with the current compilation profile.
3. Eliminate functions with the wrong number of formal parameters. If a candidate function has excess formal parameters, and each of the excess parameters has a default value, do not eliminate the function.
4. If the set is empty, fail.
5. For each actual parameter expression in sequence (left to right), perform the following:
    a. If the type of the actual parameter matches the unqualified type of the corresponding formal parameter in any function in the set, remove all functions whose corresponding parameter does not match exactly.
    b. If there is a function with a dynamically typed formal argument which is compatible with the actual parameter type, remove all functions whose corresponding parameter is not similarly compatible.
    B. If there is a defined promotion for the type of the actual parameter to the unqualified type of the formal parameter of any function, remove all functions for which this is not true from the set.
    d. If there is a valid implicit cast that converts the type of the actual parameter to the unqualified type of the formal parameter of any function, remove all functions for which this is not true from the set
    e. Fail.

6. Choose a function based on profile:
   a. If there is at least one function with a profile that exactly matches the compilation profile, discard all functions that don't exactly match.
   b. Otherwise, if there is at least one function with a wildcard profile that matches the compilation profile, determine the 'most specific' matching wildcard profile in the candidate set. Discard all functions except those with this 'most specific' wildcard profile. How 'specific' a given wildcard profile name is relative to a particular profile is determined by the profile specification.
7. If the number of functions remaining in the set is not one, then fail.

Global Variables

Global variables are declared and used as in C. Non-static variables may have a semantic associated with them. Uniform non-static variables may have their value set through the run-time API.

Use of Uninitialized Variables

It is incorrect for a program to use an uninitialized static or local variable. However, the compiler is not obligated to detect such errors, even if it would be possible to do so by compile-time data-flow analysis. The value obtained from reading an uninitialized variable is undefined. This same rule applies to the implicit use of a variable that occurs when it is returned by a top-level function. In particular, if a top-level function returns a `struct`, and some element of that `struct` is never written, then the value of that element is undefined.

Note: The language designers did not choose to define variables as being initialized to zero because that would result in a performance penalty in cases where the compiler is unable to determine if a variable is properly initialized by the programmer.

Preprocessor

Cg profiles must support the full ANSI C standard preprocessor capabilities: `#if`, `#define`, and so on. However, while `#include` must be supported the mechanism by which the file to be included is located is implementation defined.

Overview of Binding Semantics

In stream-processing architectures, data packets flow between different programmable units. On a GPU, for example, packets of vertex data flow from the application to the vertex program.

Because packets are produced by one program (the application, in this case), and consumed by another (the vertex program), there must be some mechanism for defining the interface between the two. Cg allows the user to choose between two different approaches to defining these interfaces.

The first approach is to associate a binding semantic with each element of the packet. This approach is a *bind-by-name* approach. For example, an output with the binding semantic `FOO` is fed to an input with the binding semantic `FOO`. Profiles may allow the user to define arbitrary identifiers in this ''semantic namespace'', or they may restrict the allowed identifiers to a predefined set. Often, these predefined names correspond to the names of hardware registers or API resources.

In some cases, predefined names may control non-programmable parts of the hardware. For example, vertex programs normally compute a position that is fed to the rasterizer, and this position is stored in an output with the binding semantic `POSITION`.

For any profile, there are two namespaces for predefined binding semantics — the namespace used for `in` variables and the namespace used for `out` variables. The primary implication of having two namespaces is that the binding semantic cannot be used to implicitly specify whether a variable is `in` or `out`.

The second approach to defining data packets is to describe the data that is present in a packet and allow the compiler to decide how to store it. In Cg, the user can describe the contents of a data packet by placing all of its contents into a `struct`. When a `struct` is used in this manner, we refer to it as a *connector*. The two approaches are not mutually exclusive, as is discussed later. The connector approach allows the user to rely on a combination of user-specified semantic bindings and compiler-determined bindings.

Binding Semantics

A binding semantic may be associated with an input to a top-level function or a global variable in one of three ways:

- The binding semantic is specified in the formal parameter declaration for the function. The syntax for formal parameters to a function is:

```
[const] [in | out | inout] <type> <identifier> [: <binding-semantic>] [=
```

- If the formal parameter is a `struct`, the binding semantic may be specified with an element of the `struct` when the `struct` is defined:

```
struct <struct-tag> {
    <type> <identifier>[ : <binding-semantic>];
      ...
};
```

- If the input to the function is implicit (a non-static global variable that is read by the function), the binding semantic may be specified when the non-static global variable is declared:

```
[varying [in | out]] <type> <identifier> [ : <binding-semantic>];
```

  If the non-static global variable is a `struct`, the binding semantic may be specified when the `struct` is defined, as described in the second bullet above.

- A binding semantic may be associated with the output of a top-level function in a similar manner:

```
<type> <identifier> ( <parameter-list> ) [: <binding-semantic>]
{
        :
```

Another method available for specifying a semantic for an output value is to return a `struct`, and to specify the binding *semantic* (s) with elements of the `struct` when the `struct` is defined. In addition, if the output is a formal parameter, then the binding semantic may be specified using the same approach used to specify binding semantics for inputs.

Aliasing of Semantics

Semantics must honor a copy-on-input and copy-on-output model. Thus, if the same input binding semantic is used for two different variables, those variables are initialized with the same value, but the variables are not aliased thereafter. Output aliasing is illegal, but implementations are not required to detect it. If the compiler does not issue an error on a program that aliases output binding semantics, the results are undefined.

Additional Details for Binding Semantics

The following are somewhat redundant, but provide extra clarity:

- Semantic names are case-insensitive.
- Semantics attached to parameters to non-main functions are ignored.
- Input semantics may be aliased by multiple variables.
- Output semantics may not be aliased.

Using a Structure to Define Binding Semantics (Connectors)

Cg profiles may optionally allow the user to avoid the requirement that a binding semantic be specified for every non-uniform input (or output) variable to a top-level program. To avoid this requirement, all the non-uniform variables should be included within a single `struct`. The compiler automatically allocates the elements of this structure to hardware resources in a manner that allows any program that returns this `struct` to interoperate with any program that uses this `struct` as an input.

It is not *required* that all non-uniform inputs be included within a single struct in order to omit binding semantics. Binding semantics may be omitted from any input or output, and the compiler performs automatic allocation of that input or output to a hardware resource. However, to guarantee interoperability of one program's output with another program's input when automatic binding is performed, it is necessary to put all of the variables in a single `struct`.

It is permissible to explicitly specify a binding semantic for some elements of the `struct`, but not others. The compiler's automatic allocation must honor these explicit bindings. The allowed set of explicitly specified binding semantics is defined by the allocation-rule identifier. The most common use of this capability is to bind variables to hardware registers that write to, or read from, non-programmable parts of the hardware. For example, in a typical vertex-program profile, the output `struct` would contain an element with an explicitly specified POSITION semantic. This element is used to control the hardware rasterizer.

Defining Binding Semantics via an external API

It may be possible to define binding semantics on inputs and outputs by using an external API that manipulates the programs environment. The Cg Runtime API is such an API that allows this, and others may exist.

How Programs Receive and Return Data

A program is a non-static function that has been designated as the main entry point at compilation time. The varying inputs to the program come from this top-level function's varying `in` parameters, and any global varying variables that do not have an `out` modifier. The uniform inputs to the program come from the top-level function's uniform `in` parameters and from any non-static global variables that are referenced by the top-level function or by any functions that it calls. The output of the program comes from the return value of the function (which is always implicitly varying), from any `out` parameters, which must also be varying, and from any `varying out` global variables that are written by the program.

Parameters to a program of type `sampler*` are implicitly `const`.

Statements and Expressions

Statements are expressed just as in C, unless an exception is stated elsewhere in this document. Additionally,

- `if`, `while`, and `for` require bool expressions in the appropriate places.
- Assignment is performed using `=`. The assignment operator returns a value, just as in C, so assignments may be chained.
- The new `discard` statement terminates execution of the program for the current data element (such as the current vertex or current fragment) and suppresses its output. Vertex profiles may choose to omit support for `discard`.

Minimum Requirements for if, while, for

The minimum requirements are as follows:

- All profiles should support `if`, but such support is not strictly required for older hardware.
- All profiles should support `for` and `while` loops if the number of loop iterations can be determined at compile time. "Can be determined at compile time" is defined as follows: The loop-iteration expressions can be evaluated at compile time by use of intra-procedural constant propagation and folding, where the variables through which constant values are propagated do not appear as lvalues within any kind of control statement (`if`, `for`, or `while`) or `?:` construct. Profiles may choose to support more general constant propagation techniques, but such support is not required.

- Profiles may optionally support fully general `for` and `while` loops.

New Vector Operators

These new operators are defined for vector types:

- Vector construction operator: *typeID*`(...)`:

  This operator builds a vector from multiple scalars or shorter vectors:

  - `float4(scalar, scalar, scalar, scalar)`
  - `float4(float3, scalar)`

- Matrix construction operator: *typeID*`(...)`:

  This operator builds a matrix from multiple rows.

  Each row may be specified either as multiple scalars or as any combination of scalars and vectors with the appropriate size, e.g.

```
float3x3(1, 2, 3, 4, 5, 6, 7, 8, 9)
float3x3(float3, float3, float3)
float3x3(1, float2, float3, 1, 1, 1)
```

- Vector swizzle operator: (`.`)

```
a = b.xxyz; // A swizzle operator example
```

  - At least one swizzle character must follow the operator.
  - There are three sets of swizzle characters and they may not be mixed: Set one is `xyzw = 0123`, set two is `rgba = 0123`, and set three is `stpq = 0123`.
  - The vector swizzle operator may only be applied to vectors or to scalars.
  - Applying the vector swizzle operator to a scalar gives the same result as applying the operator to a vector of length one. Thus, `myscalar.xxx` and `float3(myscalar, myscalar, myscalar)` yield the same value.
  - If only one swizzle character is specified, the result is a scalar not a vector of length one. Therefore, the expression `b.y` returns a scalar.
  - Care is required when swizzling a constant scalar because of ambiguity in the use of the decimal point character. For example, to create a three-vector from a scalar, use one of the following: `(1).xxx` or `1..xxx` or `1.0.xxx` or `1.0f.xxx`
  - The size of the returned vector is determined by the number of swizzle characters. Therefore, the size of the result may be larger or smaller than the size of the original vector. For example, `float2(0,1).xxyy` and `float4(0,0,1,1)` yields the same result.

- Matrix swizzle operator:

  For any matrix type of the form '<type><rows>x<columns>', the notation: '<matrixObject>._m<row><col>[_m<row><col>][...]' can be used to access individual matrix elements (in the case of only one <row>,<col> pair) or to construct vectors from elements of a matrix (in the case of more than one <row>,<col> pair). The row and column numbers are zero-based.

  For example:

```
float4x4 myMatrix;
float    myFloatScalar;
float4   myFloatVec4;
```

```
// Set myFloatScalar to myMatrix[3][2]
myFloatScalar = myMatrix._m32;

// Assign the main diagonal of myMatrix to myFloatVec4
myFloatVec4 = myMatrix._m00_m11_m22_m33;
```

For compatibility with the D3DMatrix data type, Cg also allows one-based swizzles, using a form with the `m` omitted after the _: '<matrixObject>._<row><col>[_<row><col>][...]' In this form, the indexes for <row> and <col> are one-based, rather than the C standard zero-based. So, the two forms are functionally equivalent:

```
float4x4 myMatrix;
float4   myVec;

// These two statements are functionally equivalent:
myVec = myMatrix._m00_m23_m11_m31;
myVec = myMatrix._11_34_22_42;
```

Because of the confusion that can be caused by the one-based indexing, its use is strongly discouraged. Also one-based indexing and zero-based indexing cannot be mixed in a single swizzle

The matrix swizzles may only be applied to matrices. When multiple components are extracted from a matrix using a swizzle, the result is an appropriately sized vector. When a swizzle is used to extract a single component from a matrix, the result is a scalar.

- The write-mask operator: `(.)` It can only be applied to an lvalue that is a vector or matrix. It allows assignment to particular elements of a vector or matrix, leaving other elements unchanged. It looks exactly like a swizzle, with the additional restriction that a component cannot be repeated.

Arithmetic Precision and Range

Some hardware may not conform exactly to IEEE arithmetic rules. Fixed-point data types do not have IEEE-defined rules.

Optimizations are permitted to produce slightly different results than unoptimized code. Constant folding must be done with approximately the correct precision and range, but is not required to produce bit-exact results. It is recommended that compilers provide an option either to forbid these optimizations or to guarantee that they are made in bit-exact fashion.

Operator Precedence

Cg uses the same operator precedence as C for operators that are common between the two languages.

The swizzle and write-mask operators `(.)` have the same precedence as the structure member operator `(.)` and the array index operator `[]`.

Operator Enhancements

The standard C arithmetic operators (`+`, `-`, `*`, `/`, `%`, `unary -`) are extended to support vectors and matrices. Sizes of vectors and matrices must be appropriately matched, according to standard mathematical rules. Scalar-to-vector promotion, as described earlier, allows relaxation of these rules.

M[n][m]
Matrix with `n` rows and `m` columns

V[n]
Vector with `n` elements

-V[n] -> V[n]
Unary vector negate

−M[n] −> M[n]
Unary matrix negate

V[n] * V[n] −> V[n]
Componentwise *

V[n] / V[n] −> V[n]
Componentwise /

V[n] % V[n] −> V[n]
Componentwise %

V[n] + V[n] −> V[n]
Componentwise +

V[n] − V[n] −> V[n]
Componentwise −

M[n][m] * M[n][m] −> M[n][m]
Componentwise *

M[n][m] / M[n][m] −> M[n][m]
Componentwise /

M[n][m] % M[n][m] −> M[n][m]
Componentwise %

M[n][m] + M[n][m] −> M[n][m]
Componentwise +

M[n][m] − M[n][m] −> M[n][m]
Componentwise −

Operators

Boolean

```
&&  ||  !
```

Boolean operators may be applied to `bool` packed bool vectors, in which case they are applied in elementwise fashion to produce a result vector of the same size. Each operand must be a `bool` vector of the same size.

Both sides of && and || are always evaluated; there is no short-circuiting as there is in C.

Comparisons

```
<  >  <=  >=  !=  ==
```

Comparison operators may be applied to numeric vectors. Both operands must be vectors of the same size. The comparison operation is performed in elementwise fashion to produce a `bool` vector of the same size.

Comparison operators may also be applied to `bool` vectors. For the purpose of relational comparisons, `true` is treated as one and `false` is treated as zero. The comparison operation is performed in elementwise fashion to produce a `bool` vector of the same size.

Comparison operators may also be applied to numeric or bool scalars.

Arithmetic

```
+  -  *  /  %  ++  --  unary-  unary+
```

The arithmetic operator % is the remainder operator, as in C. It may only be applied to two operands of `cint` or `int` types.

When / or % is used with `cint` or `int` operands, C rules for integer / and % apply.

The C operators that combine assignment with arithmetic operations (such as `+=`) are also supported when the corresponding arithmetic operator is supported by Cg.

Conditional Operator

```
?:
```

If the first operand is of type `bool`, one of the following must hold for the second and third operands:

- Both operands have compatible structure types.
- Both operands are scalars with numeric or `bool` type.
- Both operands are vectors with numeric or `bool` type, where the two vectors are of the same size, which is less than or equal to four.

If the first operand is a packed vector of `bool`, then the conditional selection is performed on an element-wise basis. Both the second and third operands must be numeric vectors of the same size as the first operand.

Unlike C, side effects in the expressions in the second and third operands are always executed, regardless of the condition.

Miscellaneous Operators

```
(typecast)   ,
```

Cg supports C's typecast and comma operators.

Reserved Words

The following are currently used reserved words in Cg. A '*' indicates that the reserved word is case-insensitive.

__[anything] (i.e. any identifier with two underscores as a prefix)
asm*
asm_fragment
auto
bool
break
case
catch
char
class
column_major
compile
const
const_cast
continue
decl*
default
delete
discard
do
double
dword*
dynamic_cast
else
emit

enum
explicit
extern
false
fixed
float*
for
friend
get
goto
half
if
in
inline
inout
int
interface
long
matrix*
mutable
namespace
new
operator
out
packed
pass*
pixelfragment*
pixelshader*
private
protected
public
register
reinterpret_cast
return
row_major
sampler
sampler_state
sampler1D
sampler2D
sampler3D
samplerCUBE
shared
short
signed
sizeof
static
static_cast
string*
struct
switch
technique*
template

texture*
texture1D
texture2D
texture3D
textureCUBE
textureRECT
this
throw
true
try
typedef
typeid
typename
uniform
union
unsigned
using
vector*
vertexfragment*
vertexshader*
virtual
void
volatile
while

Cg Standard Library Functions

Cg provides a set of built-in functions and structures to simplify GPU programming. These functions are similar in spirit to the C standard library functions, providing a convenient set of common functions.

The Cg Standard Library is documented in ''spec_stdlib.txt''.

VERTEX PROGRAM PROFILES

A few features of the Cg language that are specific to vertex program profiles are required to be implemented in the same manner for all vertex program profiles.

Mandatory Computation of Position Output

Vertex program profiles may (and typically do) require that the program compute a position output. This homogeneous clip-space position is used by the hardware rasterizer and must be stored in a program output with an output binding semantic of `POSITION` (or `HPOS` for backward compatibility).

Position Invariance

In many graphics APIs, the user can choose between two different approaches to specifying per-vertex computations: use a built-in configurable ''fixed-function'' pipeline or specify a user-written vertex program. If the user wishes to mix these two approaches, it is sometimes desirable to guarantee that the position computed by the first approach is bit-identical to the position computed by the second approach. This ''position invariance'' is particularly important for multipass rendering.

Support for position invariance is optional in Cg vertex profiles, but for those vertex profiles that support it, the following rules apply:

- Position invariance with respect to the fixed function pipeline is guaranteed if two conditions are met:
  - A `#pragma position_invariant <top-level-function-name>` appears before the body of the top-level function for the vertex program.
  - The vertex program computes position as follows:

```
OUT_POSITION = mul(MVP, IN_POSITION)
```

where:

OUT_POSITION
: is a variable (or structure element) of type `float4` with an output binding semantic of `POSITION` or `HPOS`.

IN_POSITION
: is a variable (or structure element) of type `float4` with an input binding semantic of `POSITION`.

MVP
: is a uniform variable (or structure element) of type `float4x4` with an input binding semantic that causes it to track the fixed-function modelview-projection matrix. (The name of this binding semantic is currently profile-specific — for OpenGL profiles, the semantic `state.matrix.mvp` is recommended).

- If the first condition is met but not the second, the compiler is encouraged to issue a warning.
- Implementations may choose to recognize more general versions of the second condition (such as the variables being copy propagated from the original inputs and outputs), but this additional generality is not required.

Binding Semantics for Outputs

As shown in Table 10, there are two output binding semantics for vertex program profiles:

Table 10 Vertex Output Binding Semantics

Name	Meaning	Type	Default Value
POSITION	Homogeneous clip-space position; fed to rasterizer.	float4	Undefined
PSIZE	Point size	float	Undefined

Profiles may define additional output binding semantics with specific behaviors, and these definitions are expected to be consistent across commonly used profiles.

FRAGMENT PROGRAM PROFILES

A few features of the Cg language that are specific to fragment program profiles are required to be implemented in the same manner for all fragment program profiles.

Binding semantics for outputs

As shown in Table 11, there are three output binding semantics for fragment program profiles:

Table 11 Fragment Output Binding Semantics

Name	Meaning	Type	Default Value
COLOR	RGBA output color	float4	Undefined
COLOR0	Same as COLOR		
DEPTH	Fragment depth value (in range [0,1])	float	Interpolated depth from rasterizer (in range [0,1])

Profiles may define additional output binding semantics with specific behaviors, and these definitions are expected to be consistent across commonly used profiles.

If a program desires an output color alpha of 1.0, it should explicitly write a value of 1.0 to the `W` component of the `COLOR` output. The language does *not* define a default value for this output.

Note: If the target hardware uses a default value for this output, the compiler may choose to optimize away an explicit write specified by the user if it matches the default hardware value. Such defaults are not exposed in the language.)

In contrast, the language does define a default value for the DEPTH output. This default value is the interpolated depth obtained from the rasterizer. Semantically, this default value is copied to the output at the beginning of the execution of the fragment program.

As discussed earlier, when a binding semantic is applied to an output, the type of the output variable is not required to match the type of the binding semantic. For example, the following is legal, although not recommended:

```
struct myfragoutput {
    float2 mycolor : COLOR;
}
```

In such cases, the variable is implicitly copied (with a typecast) to the semantic upon program completion. If the variable's vector size is shorter than the semantic's vector size, the larger-numbered components of the semantic receive their default values if applicable, and otherwise are undefined. In the case above, the R and G components of the output color are obtained from mycolor, while the B and A components of the color are undefined.