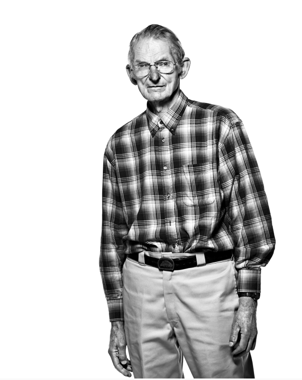
6 Algorithmic Journeys with Concepts

Taras Shevchenko

Rails Reactor / Giphy

The Software Industry is Not Industrialized

Software components (routines), to be widely applicable to different machines and users, should be available in families arranged according to precision, robustness, generality and time-space performance.



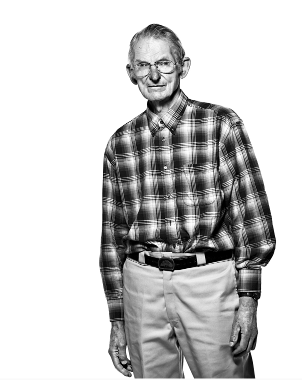
A Familiar Example. Douglas McIlroy about sin

Dimensions along which we wish to have variablity:

1. precision, for which perhaps ten different approximating functionsmight suffice
2. floating vs fixed computation
3. argument ranges $[0, \pi/2]$, $[0, 2pi]$, also $[-\pi/2, pi/2]$, $[-\pi, \pi]$, $[-big, +big]$
4. robustness - ranging from no argument validation through signaling of complete loss of significance, to signaling of specified range violations

Douglas McIlroy

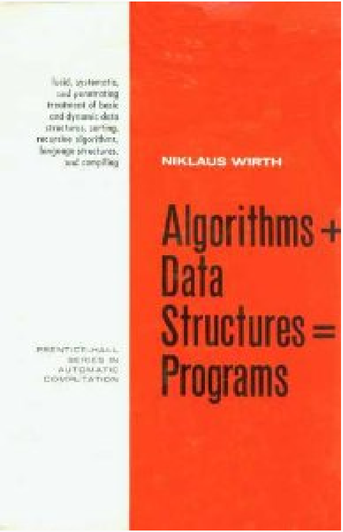
1. Choices
    1.1 precision
    1.2 robustness
    1.3 generality
    1.4 generality
    1.5 algorithm
    1.6 interfaces and error-handling
2. Application Areas
    2.1 numerical approximation
    2.2 I/O
    2.3 2d and 3d geometry
    2.4 text processing
    2.5 storage management



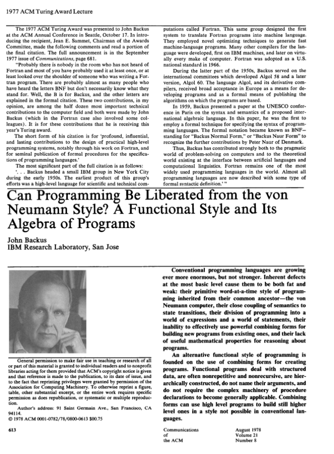
Donald Knuth

1. Fundamental Algorithms
2. Seminumerical Algorithms
3. Sorting and Searching
4. Combinatorial Algorithms
5. Syntactic Algorithms
6. The Theory of Context-free Languages
7. Compiler Techniques

Niklaus Wirth



John Backus



1977 ACM Turing Award Lecture

The 1977 ACM Turing Award was presented to John Backus at the ACM Annual Conference in Seattle, October 17. In introducing the recipient, Jean E. Sammet, Chairman of the Awards Committee, made the following comments and read a portion of the final citation. The full announcement is in the September 1977 issue of *Communications*, page 681.

"Probably there is nobody in the room who has not heard of Fortran and most of you have probably used it at least once, or at least looked over the shoulder of someone who was writing a Fortran program. There are probably almost as many people who have heard the letters BNF but don't necessarily know what they stand for. Well, the B is for Backus, and the other letters are explained in the formal citation. These two contributions, in my opinion, are among the half dozen most important technical contributions to the computer field and both were made by John Backus (which in the Fortran case also involved some colleagues). It is for these contributions that he is receiving this year's Turing award.

The short form of his citation is for 'profound, influential, and lasting contributions to the design of practical high-level programming systems, notably through his work on Fortran, and for seminal publication of formal procedures for the specifications of programming languages.'

The most significant part of the full citation is as follows:

'. . . Backus headed a small IBM group in New York City during the early 1950s. The earliest product of this group's efforts was a high-level language for scientific and technical computations called Fortran. This same group designed the first system to translate Fortran programs into machine language. They employed novel optimizing techniques to generate fast machine-language programs. Many other compilers for the language were developed, first on IBM machines, and later on virtually every make of computer. Fortran was adopted as a U.S. national standard in 1966.

During the latter part of the 1950s, Backus served on the international committees which developed Algol 58 and a later version, Algol 60. The language Algol, and its derivative compilers, received broad acceptance in Europe as a means for developing programs and as a formal means of publishing the algorithms on which the programs are based.

In 1959, Backus presented a paper at the UNESCO conference in Paris on the syntax and semantics of a proposed international algebraic language. In this paper, he was the first to employ a formal technique for specifying the syntax of programming languages. The formal notation became known as BNF—standing for "Backus Normal Form," or "Backus Naur Form" to recognize the further contributions by Peter Naur of Denmark.

Thus, Backus has contributed strongly both to the pragmatic world of problem-solving on computers and to the theoretical world existing at the interface between artificial languages and computational linguistics. Fortran remains one of the most widely used programming languages in the world. Almost all programming languages are now described with some type of formal syntactic definition.' "

Can Programming Be Liberated from the von Neumann Style? A Functional Style and Its Algebra of Programs

John Backus
IBM Research Laboratory, San Jose

Conventional programming languages are growing ever more enormous, but not stronger. Inherent defects at the most basic level cause them to be both fat and weak: their primitive word-at-a-time style of programming inherited from their common ancestor—the von Neumann computer, their close coupling of semantics to state transitions, their division of programming into a world of expressions and a world of statements, their inability to effectively use powerful combining forms for building new programs from existing ones, and their lack of useful mathematical properties for reasoning about programs.

An alternative functional style of programming is founded on the use of combining forms for creating programs. Functional programs deal with structured data, are often nonrepetitive and nonrecursive, are hierarchically constructed, do not name their arguments, and do not require the complex machinery of procedure declarations to become generally applicable. Combining forms can use high level programs to build still higher level ones in a style not possible in conventional languages.

General permission to make fair use in teaching or research of all or part of this material is granted to individual readers and to nonprofit libraries acting for them provided that ACM's copyright notice is given and that reference is made to the publication, to its date of issue, and to the fact that reprinting privileges were granted by permission of the Association for Computing Machinery. To otherwise reprint a figure, table, other substantial excerpt, or the entire work requires specific permission as does republication, or systematic or multiple reproduction.

Author's address: 91 Saint Germain Ave., San Francisco, CA 94114.
© 1978 ACM 0001-0782/78/0800-0613 $00.75

613

Communications of the ACM — August 1978, Volume 21, Number 8

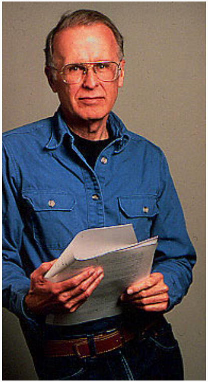


Figure 1: We need a few functional forms

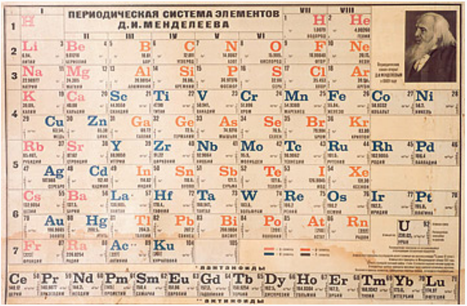
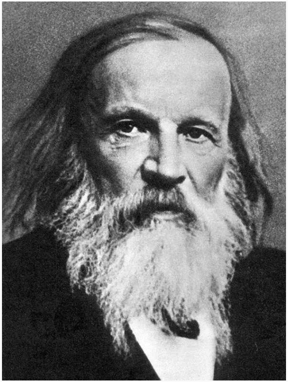


Figure 2: A Russian periodic table based on Dmitri Mendeleyev's original table of 1869.



Carl Linnaeus

Species Plantarum lists every species of plant known at the time, classified into genera. It is the first work to consistently apply binomial names and was the starting point for the naming of plants.

Euclid

1. Definitions
2. Postulates
3. Common notions

Common Notions

1. Things which are equal to the same thing are also equal to one other.
2. If equals be added to equals, the wholes are equal.
3. If equals be subructed from equals, the remainders are equal.
4. Things which coincide with one another are equal to one another.
5. The whole is greater than the part.

Basic idea

The essence of generic programming lies in the idea of concepts. A concept is a way of describing a family of related object types.

Natural Science	**Mathematics**	**Programming**	**Programming Examples**
genus	theory	concept	Integral, Character
species	model	type or class	uint8_t, char
individual	element	instance	01000001(65, 'A')

Definitions

1. Datum
2. Value
3. Value type
4. Object
5. Object type

Datum

Definition
A **datum** is a sequence of bits.

Example
01000001 is an example of a datum.

Value

Definition
A **value is** a **datum** together with its interpretation.

Example
The **datum** 01000001 might have the interpretation of the integer 65, or the character “A”.

Explanation
Every **value** must be associated with a **datum** in memory; there is no way to refer to disembodied **values** in modern programming languages.

Value type

Definition
A **value type** is a set of values sharing a common interpretation.

Object

Definition
An **object** is a collection of bits in memory that contain a **value** of a given **value type**.

Explanation
An **object** is immutable if the value never changes, and mutable otherwise. An object is unrestricted if it can contain any **value** of its **value type**.

Object type

Definition
An **object type** is a uniform method of storing and retrieving **values** of a given **value type** from a particular **object** when given its address.

Programming with concepts

Semiregular

Operation

1. Copy construction
2. Assignment
3. Destruction

Semantic

$$\forall a\ \forall b\ \forall c : T\ a(b) \implies (b = c \implies a = c)$$

$$\forall a\ \forall b\ \forall c : a \leftarrow b \implies (b = c \implies a = c)$$

$$\forall f \in RegularFunction : a = b \implies f(a) = f(b)$$

Semiregular

```
template<typename T>
concept semiregular = std::is_default_constructible<T>::value &&
                      std::is_copy_constructible<T>::value &&
                      std::is_copy_assignable<T>::value &&
                      std::is_destructible<T>::value
```

Regular type

Operation

1. Copy construction
2. Assignment
3. Equality
4. Destruction

Semantic

$$\forall a\ \forall b\ \forall c : T\ a(b) \implies (b = c \implies a = c)$$

$$\forall a\ \forall b\ \forall c : a \leftarrow b \implies (b = c \implies a = c)$$

$$\forall f \in RegularFunction : a = b \implies f(a) = f(b)$$

Regular

```
template<typename T>
concept semiregular = semiregular<T> && is_equality_comparable<T>::value;
```

Relation

$$\begin{aligned} \textit{FunctionalProcedure}(F) \triangleq\ & F \textit{ is a regular procedure defined on regular} \\ & \textit{types: replacing its inputs with equal objects} \\ & \textit{results in equal output objects.} \end{aligned}$$

$$\begin{aligned} \textit{HomogeneousFunction}(F) \triangleq\ & \textit{FunctionalProcedure}(F) \wedge \textit{Arity}(F) > 0 \\ & \wedge (\forall i, j \in \mathbb{N})(i, j < \textit{Arity}(F)) \implies (\textit{InputType}(F, i) = \textit{InputType}(F, j)) \\ & \wedge \textit{Domain} : \textit{HomogeneousFunction} \to \textit{Regular} \\ & F \implies \textit{InputType}(F, 0) \end{aligned}$$

Regular

```
template<typename F, typename... T>
concept functional_procedure = (regular<typename std::invoke_result<F, T
    ...>::type> || std::is_same<typename std::invoke_result<F, T...>::
    type, void>::value) && is_regular<T...>::value;

template<typename F, typename... T>
concept homogeneous_function = functional_procedure<F, T...> && sizeof
    ...(T) > 0 && all_same<T...>::value && all_regular<T>;
```

$$Predicate(P) \triangleq FunctionalProcedure(F) \wedge Codomain(P) = bool$$

$$HomogeneousPredicate(P) \triangleq Predicate(P) \wedge HomogeneousFunction(P)$$

$$Relation(R) \triangleq HomogeneousPredicate(R) \wedge Arity(R) = 2$$

Relation

```cpp
template<typename F, typename... T>
concept predicate = functional_procedure<F, T...> && std::is_same<
    codomain_t<F, T...>, bool>::value;

template<typename F, typename... T>
concept homogeneous_predicate = predicate<F, T...> &&
    homogeneous_function<T...>;

template<typename R, typename T>
concept relation = predicate<R, T, T>;
```

Totally Ordered

$$\begin{aligned} &property(R : Relation) \\ &transitive : R \\ r \mapsto (\forall a, b, c \in Domain(R))(r(a, b) \wedge r(b, c) &\implies r(a, c)) \end{aligned}$$

$$\begin{aligned} &property(R : Relation) \\ &total\_ordering : R \\ r \mapsto transitive(r) \;\wedge (\forall a, b \in Domain(R))\; &exactly\ one\ of\ following\ holds : \\ &r(a, b),\ r(b, a),\ or\ a = b \end{aligned}$$

$$TotallyOrdered(T) \triangleq Regular(T) \wedge <: T\ x\ T \rightarrow bool \wedge total\_ordering(<)$$

Totally Ordered

```
template<typename T>
concept totally_ordered = regular<T> && is_less_than_comprable<T>::value;
```

Concepts

$$
\begin{aligned}
\mathit{Readable}(T) \triangleq\ & \mathit{Regular}(T) \wedge \\
& \mathit{ValueType} : \mathit{Readable} \rightarrow \mathit{Regular} \wedge \\
& \mathit{source} : T \rightarrow \mathit{ValueType}(T) \wedge
\end{aligned}
$$

Concepts

$$
\begin{aligned}
Writable(T) \triangleq\ & Regular(T) \wedge \\
& ValueType : Writable \rightarrow Regular \wedge \\
& (\forall x \in T)(\forall v \in ValueType(T))\ sink(x) \leftarrow v \\
& \quad is\ a\ well-formed\ statement \\
& source : T \rightarrow ValueType(T) \wedge
\end{aligned}
$$

Concepts

$$\begin{aligned} Iterator(T) \triangleq\ & Regular(T) \wedge \\ & DistanceType : Iterator \rightarrow Integer \wedge \\ & successor : T \rightarrow T \wedge \\ & successor\ is\ not\ necessarily - regular \end{aligned}$$

Concepts

$$\mathit{ForwardIterator}(T) \triangleq \mathit{Iterator}(T) \land \mathit{regular\_unary\_function}(\mathit{successor})$$

Concepts

$$
\begin{aligned}
\mathit{BidirectionalIterator}(T) \triangleq\ & \mathit{ForwardIterator}(T) \wedge \\
& \mathit{predecessor} : T \to T \wedge \\
& \mathit{predecessor\ takes\ constant\ type} \wedge \\
& (\forall i \in T)\mathit{successor}(i)\mathit{isdefined} \implies \\
& \quad \mathit{predecessor}(\mathit{successor}(i))\ \mathit{is\ defined} \\
& \quad \mathit{and\ equals\ to}\ i \wedge \\
& (\forall i \in T)\mathit{predecessor}(i)\ \mathit{is\ defined} \implies \\
& \quad \mathit{successor}(\mathit{predecessor}(i))\ \mathit{is\ defined} \\
& \quad \mathit{and\ equals\ to}\ i
\end{aligned}
$$

Concepts

$$
\begin{aligned}
\mathit{IndexedIterator}(T) \triangleq\ & \mathit{ForwardIterator}(T) \wedge \\
& + : T \text{ x } \mathit{DifferenceType}(T) \rightarrow T \wedge \\
& - : T \text{ x } T \rightarrow \mathit{DifferenceType}(T) \wedge \\
& + \textit{ takes constant time} \\
& - \textit{ takes constant time}
\end{aligned}
$$

Concepts

$$
\begin{aligned}
\mathit{RandomAccessIterator}(T) \triangleq\ & \mathit{BidirectionalIterator}(T) \wedge \\
& \mathit{IndexedIterator}(T) \wedge \\
& \mathit{TotallyOrdered}(T) \wedge \\
& (\forall i, j \in T)\ i < j \iff i \prec j \wedge \\
& \mathit{DifferenceType} : \\
& \quad \mathit{RandomAccessIterator} \rightarrow \mathit{Integer} \wedge \\
& + : T\ x\ \mathit{DifferenceType}(T) \rightarrow T \wedge \\
& - : T\ x\ \mathit{DifferenceType}(T) \rightarrow T \wedge \\
& - : T\ x\ T \rightarrow \mathit{DifferenceType}(T) \wedge \\
& < \textit{takes constant time} \wedge \\
& - \textit{between and iterator and an integer} \\
& \quad \textit{takes constant time}
\end{aligned}
$$

Concepts

$$\begin{aligned}
\mathit{FunctionalProcedure}(F) \triangleq\ & F \textit{ is a regular procedure defined on regular} \\
& \textit{types : replacing its inputs with equal objects} \\
& \textit{results in equal output objects.}
\end{aligned}$$

$$\begin{aligned}
\mathit{UnaryFunction}(F) \triangleq\ & \mathit{FunctionalProcedure}(F) \wedge \mathit{Arity}(F) = 1 \\
& \wedge\ \mathit{Domain} : \mathit{UnaryFunction} \to \mathit{Regular} \\
& F \mapsto \mathit{InputType}(F, 0)
\end{aligned}$$

$$\begin{aligned}
\mathit{HomogeneousFunction}(F) \triangleq\ & \mathit{FunctionalProcedure}(F) \wedge \mathit{Arity}(F) > 0 \\
& \wedge (\forall i, j \in \mathbb{N})(i, j < \mathit{Arity}(F)) \implies (\mathit{InputType}(F, i) = \mathit{InputType}(F, j)) \\
& \wedge \mathit{Domain} : \mathit{HomogeneousFunction} \to \mathit{Regular} \\
& F \implies \mathit{InputType}(F, 0)
\end{aligned}$$

Concepts

$$Predicate(P) \triangleq FunctionalProcedure(F) \wedge Codomain(P) = bool$$

$$HomogeneousPredicate(P) \triangleq Predicate(P) \wedge HomogeneousFunction(P)$$

$$Relation(R) \triangleq HomogeneousPredicate(R) \wedge Arity(R) = 2$$

$$
\begin{aligned}
&\mathit{property}(R : \mathit{Relation}) \\
&\mathit{total\_ordering} : R \\
r \mapsto \mathit{transitive}(r)\ \wedge (\forall a, b \in \mathit{Domain}(R))\ &\mathit{exactly\ one\ of\ following\ holds} : \\
&r(a,b),\ r(b,a),\ or\ a = b
\end{aligned}
$$

Journey 1

1. min
2. max

Journey #1

```
int min(int x, int y) {
    if (y < x) {
        return y;
    }
    return x;
}
```

Journey #1

```
int min(int x, int y) {
    if (y < x) {
        return y;
    }
    return x;
}


double min(double x, double y) {
    if (y < x) {
        return y;
    }
    return x;
}
```

Journey #1

```
template<typename T>
T min(T x, T y) {
    if (y < x) {
        return y;
    }
    return x;
}
```

Journey #1

Dealing with large objects

```
template<typename T>
const T& min(const T& x, const T& y) {
    if (y < x) {
        return y;
    }
    return x;
}
```

Journey #1

```
template<typename T, typename P>
const T& min(const T& x, const T& y, P pred) {
    if (pred(y, x)) {
        return y;
    }
    return x;
}
```

Journey #1

```
struct employee {
    std::string full_name;
    int64_t salary;
};

void usage() {
  employee e0{"Bjarne␣Stroustrup", 9999999ll};
  employee e1{"Alex␣Stepanov", 9999999ll};
  min(e0, e1, [](const auto& x, const auto& y) {
    return x.salary < y.salary;
  }).salary += 10000ll;
}
```

Journey #1

```
template<typename T, typename P>
T& min(T& x, T& y, P pred) {
    if (pred(y, x)) {
        return y;
    }
    return x;
}
```

Journey #1

```
template<Regular T, Relation r>
const T& min(const T& x, const T& y, Relation r) {
    if (r(y, x)) { return y; }
    return x;
}

template<Regular T, Relation r>
T& min(T& x, T& y, Relation r) {
    if (r(y, x)) { return y; }
    return x;
}
```

Journey #1

```
template<TotallyOrdered T>
const T& min(const T& x, const T& y) {
    return min(x, y, std::less<T>());
}

template<TotallyOrdered T>
T& min(T& x, T& y) {
    return min(x, y, std::less<T>());
}
```

Journey #1

Journey #1

```
namespace cppcon {

template<totally_ordered T>
const T& min(const T& x, const T& y) {
    if (y < x) {
        return y;
    }
    return x;
}

template<totally_ordered T>
T& min(T& x, T& y) {
    if (y < x) {
        return y;
    }
    return x;
```

Journey #2

1. unique
2. unique_count

Journey #2

}

Journey #3

1. frequencies
2. transform_subgroups
3. squash_subgroups

Journey #4

1. split
2. transform_splits

Journey #5

1. remove_if
2. partition_semistable

More examples of concepts

1. Regular Type
2. Semiegular Type
3. Functional Procedure
4. Homogeneous Function
5. Homogeneous Predicate
6. Semiring
7. Sequence
8. Totally Ordered
9. Input Iterator
10. Forfward Iterator
11. Bidirectional Iterator

Properties

1. Associative
2. Distributive
3. Transitive
4. Semiegular Type
5. Functional Procedure

Techniques

1. Transformation-action duality
2. Operation-accumulation procedure duality
3. Memory adaptivity
4. Reduction to constrained subproblem

Conclusion

Conclusion

1. Concepts are mathematical. They are not specific to C++.
2. Know as many algorithms as you can.
3. Algorithms come in groups.
4. Transform complicated loops into well-defined algorithms.
5. Use mathematics for everything you do.
6. Don't obay mathematical convetions in programming.
7. Prefare concrete algorithms to more general.
8. Have a little Euclid, Knuth, Dijkstra in your mind and let them argue.