

C11

ISO/IEC 9899:2011

07.11.2013, Daniel Martens

Einführung

in C

Hello World!

```
#include <stdio.h>
```

```
int main(int argc, char *argv[])
```

```
{
```

```
    printf("Hello world!\n");
```

```
    return 0;
```

```
}
```

Entstehung

- 1969 - 1973 in den AT&T Bell Labs von Dennis Ritchie entwickelt
- Imperative Programmiersprache
- Beeinflusst von B, BCPL und ALGOL 68

Entstehung

1954

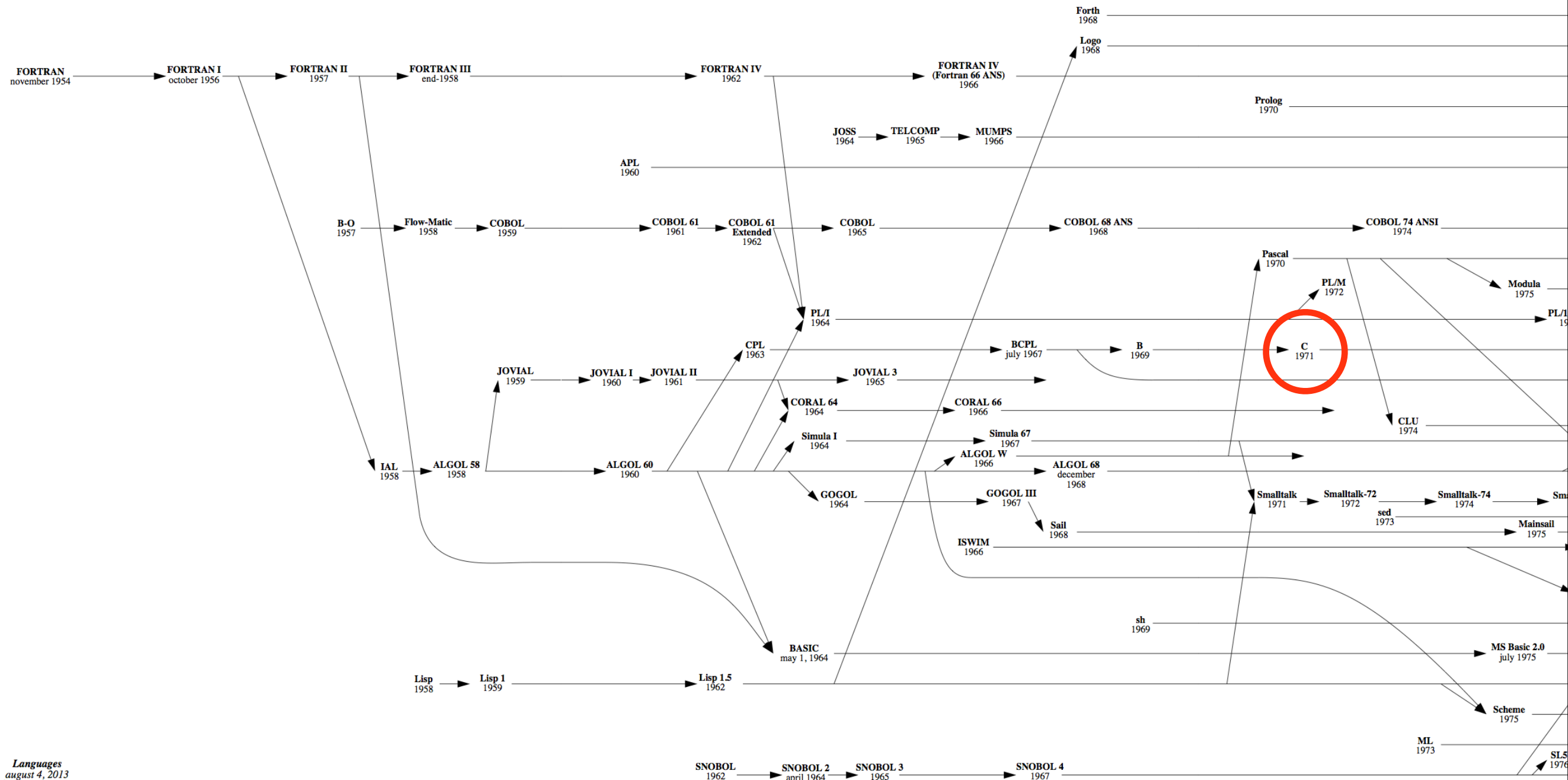
1957

1960

1965

1970

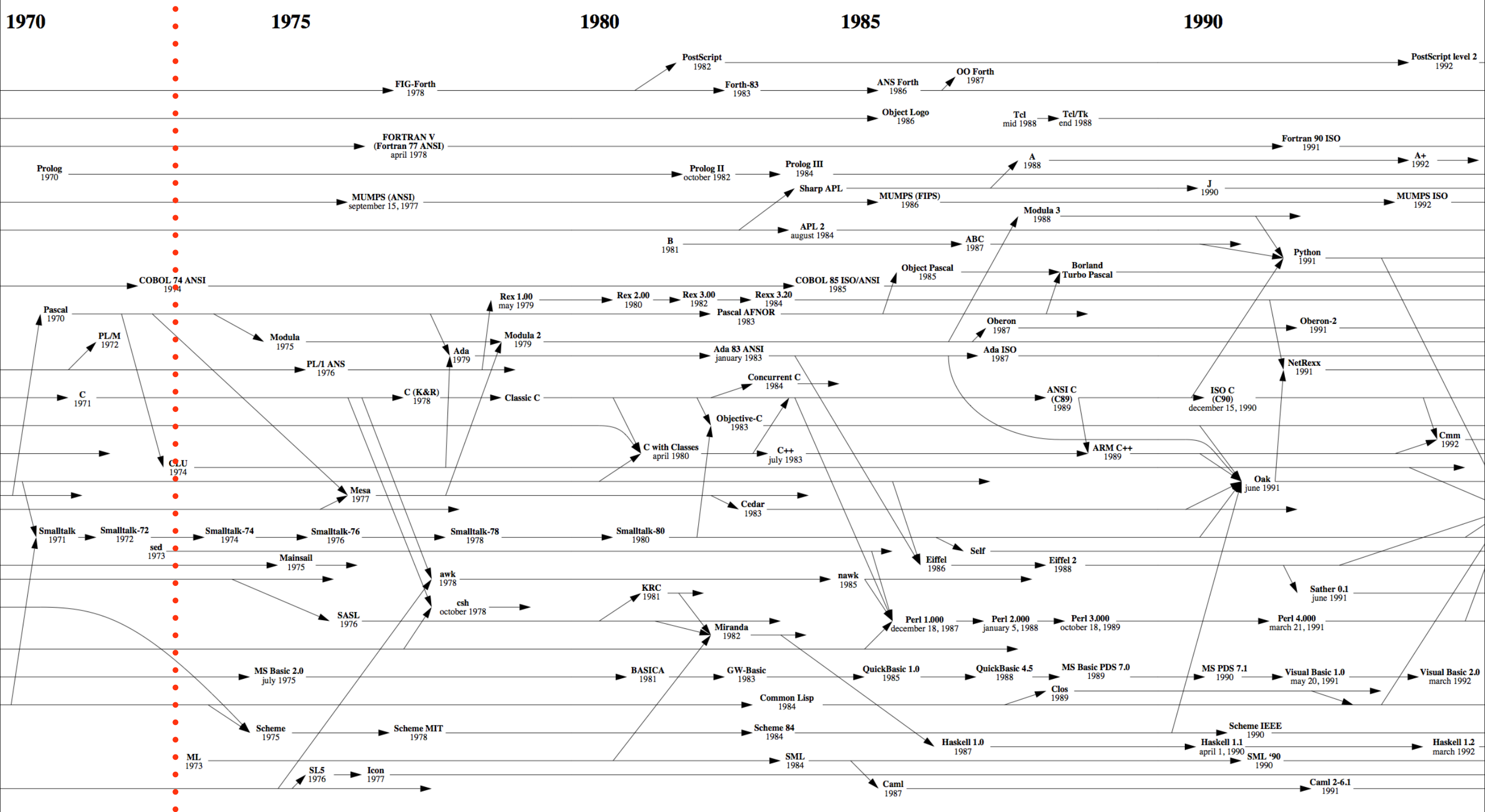
1975



Evolution

- 1973 ist C so weit ausgereift, dass Unix-Kernel, der PDP-11, in C neu geschrieben werden kann
- Erster C-Compiler ebenfalls von Dennis Ritchie entwickelt

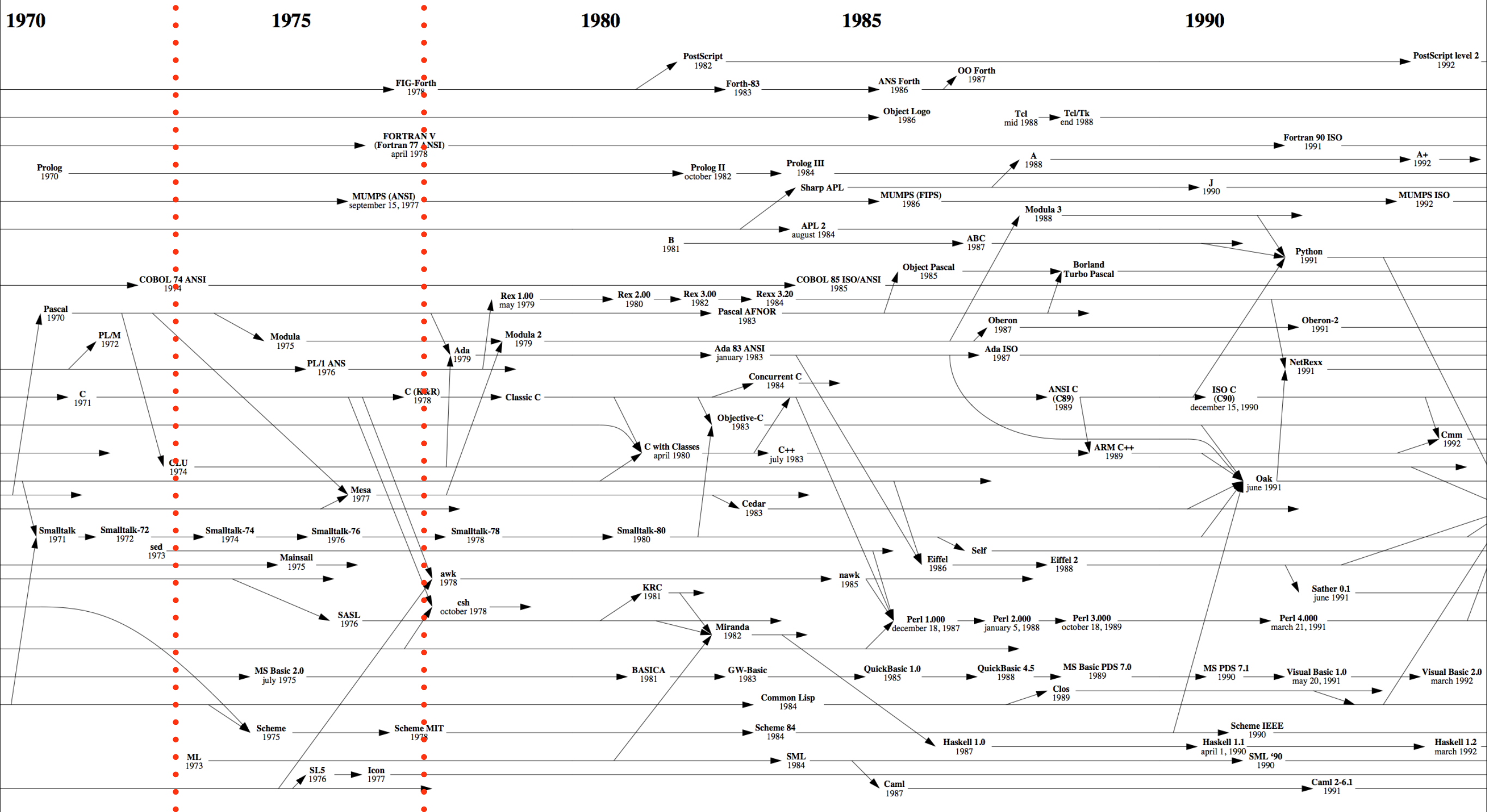
Evolution



Evolution

- 1973 ist C so weit ausgereift, dass Unix-Kernel, der PDP-11, in C neu geschrieben werden kann
- Erster C-Compiler ebenfalls von Dennis Ritchie entwickelt
- 1978 Brian W. Kernighan und Dennis Ritchie veröffentlichen erste Spezifikation („The C Programming Language“)

Evolution



Features

- Wenige Schlüsselwörter
- Standard Bibliothek
- Kompiliert zu nativem Code
- Systemnah, plattformunabhängig
- Präprozessor
- Internationaler Standard
- Compiler für nahezu jede Plattform

Anwendungsgebiete

- Systemprogrammierung
 - Betriebssysteme
- Mikrocontroller
- Eingebettete Prozessoren
- DSP Prozessoren
- Als Interpreter anderer Programmiersprachen (z.B. Java Virtual Machine)
- als Zwischencode einiger Programmiersprachen (z.B. Vala)

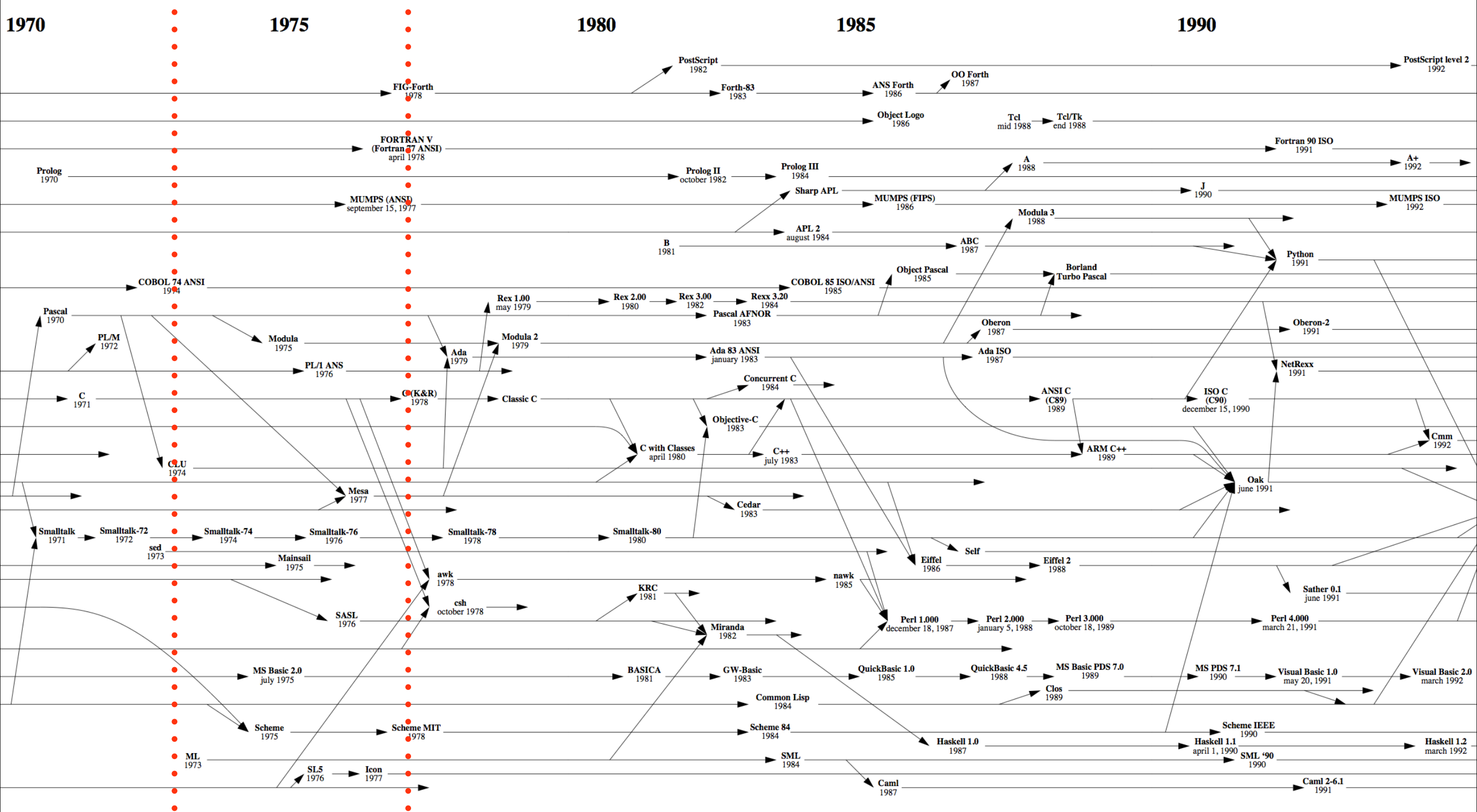
Warum C?

- Portierbarkeit
- Geschwindigkeit
- Größe
- Low-Level-Access/Hardware ansprechen
- Keine Laufzeitumgebung erforderlich
- Wird bereits lange in der Entwicklung von Betriebssystemen eingesetzt

C und andere

- Abwandlungen
 - C++, C#, Objective-C
- Auswirkungen auf
 - Java, Perl, D, PHP, Vala
- Nachteile
 - Exceptions
 - Range checking
 - Garbage collection
 - OOP

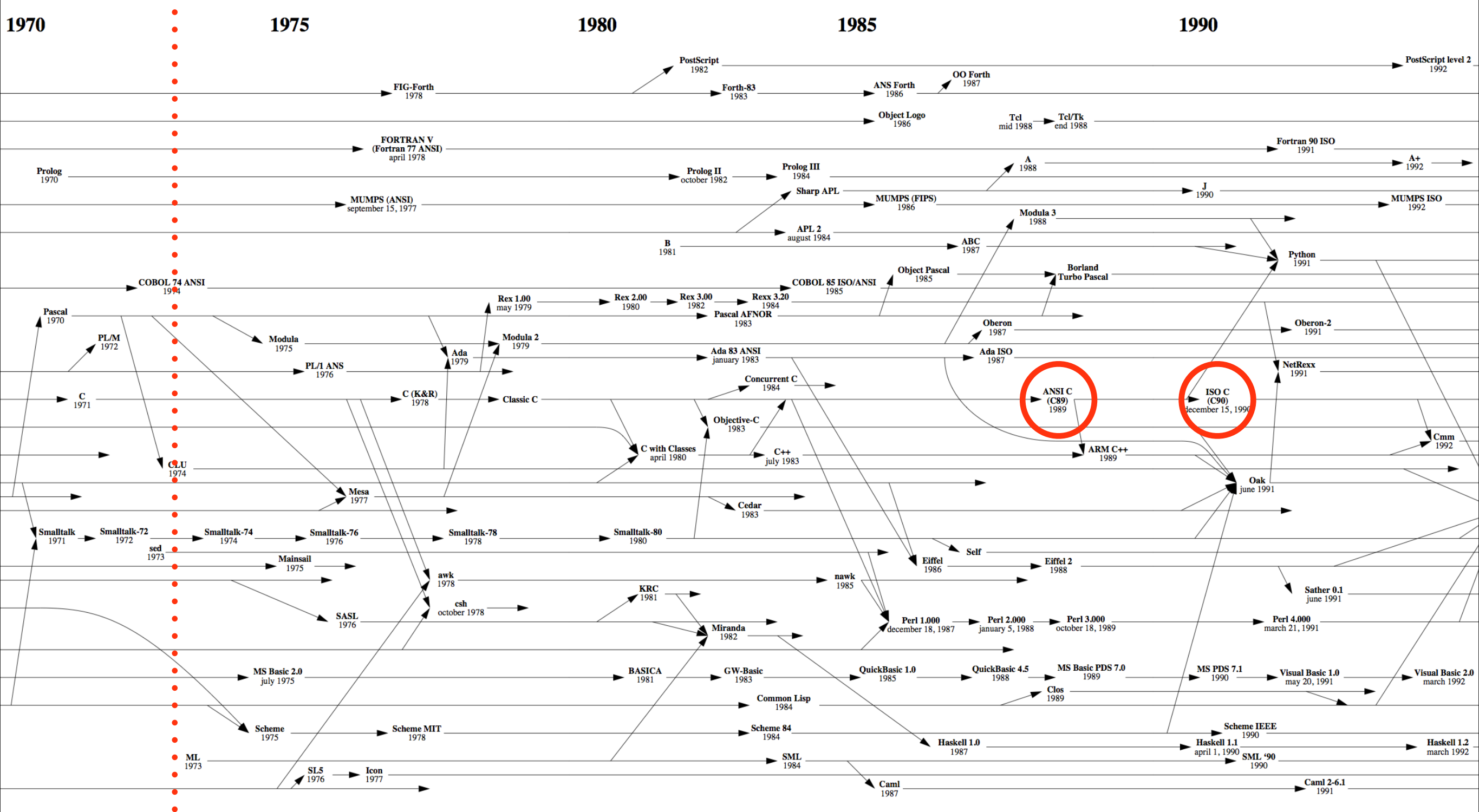
Evolution



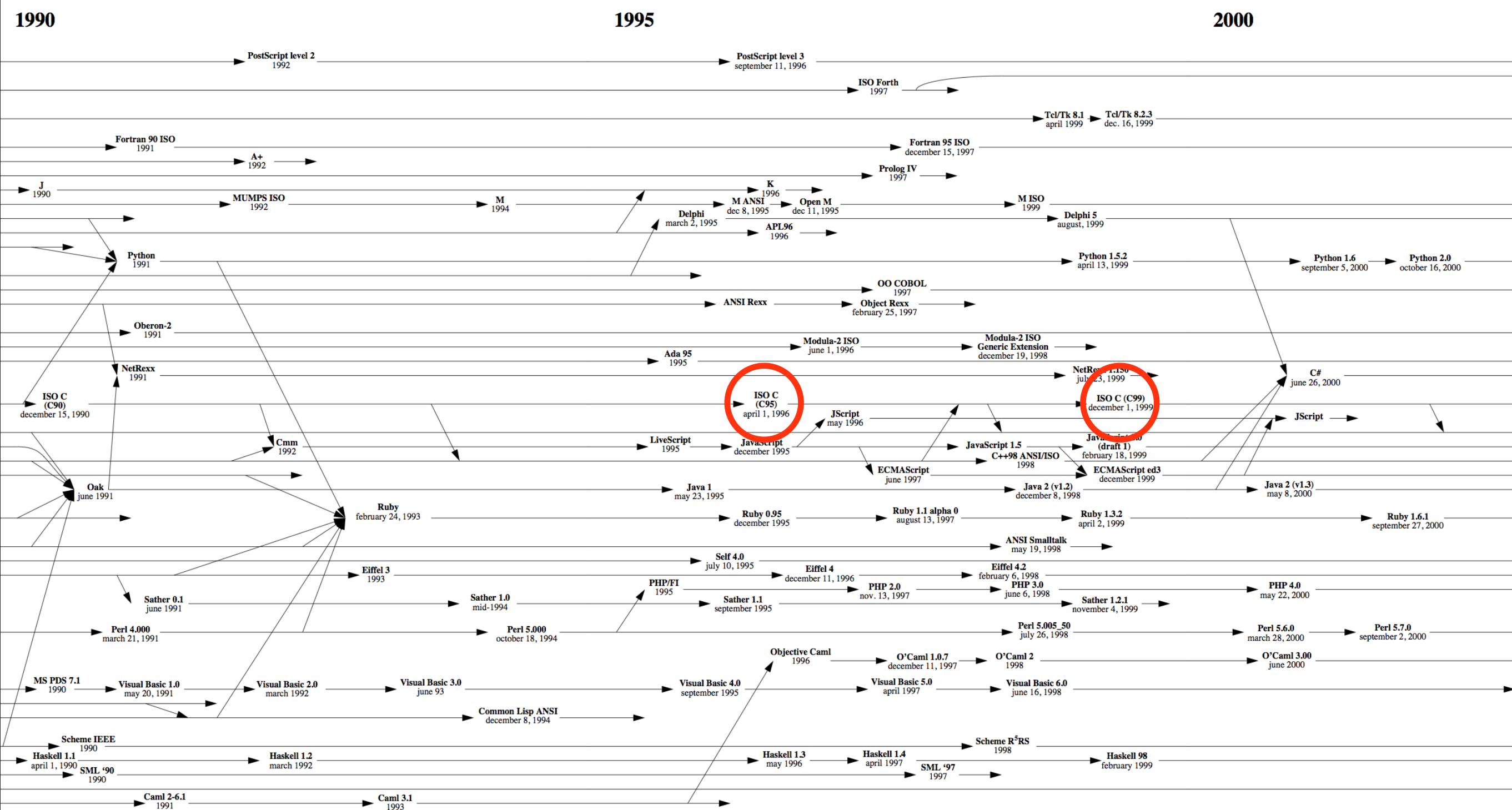
Evolution

- 1973 ist C so weit ausgereift, dass Unix-Kernel, der PDP-11, in C neu geschrieben werden kann
- Erster C-Compiler ebenfalls von Dennis Ritchie entwickelt
- 1978 Brian W. Kernighan und Dennis Ritchie veröffentlichen erste Spezifikation („The C Programming Language“)
- Standards
 - C89/C90, C95, C99, C11

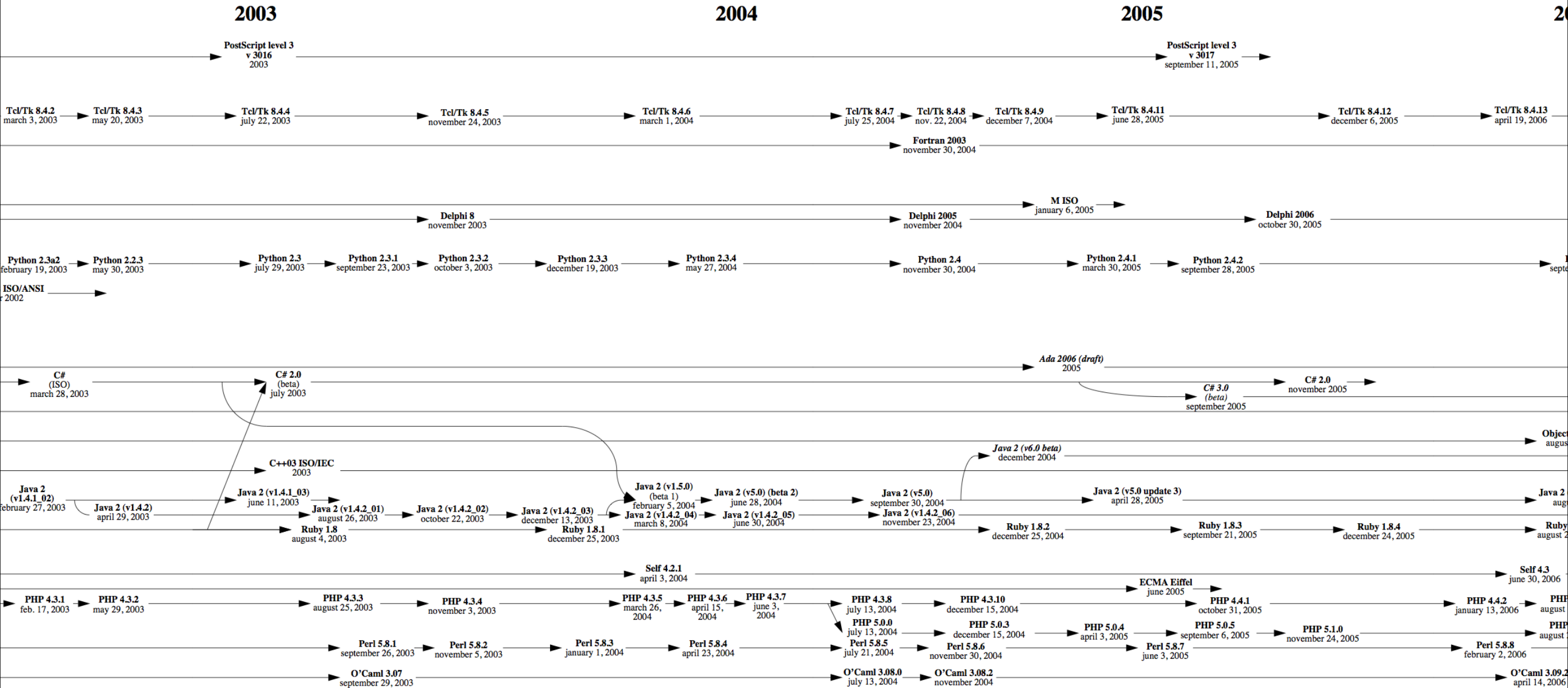
Evolution



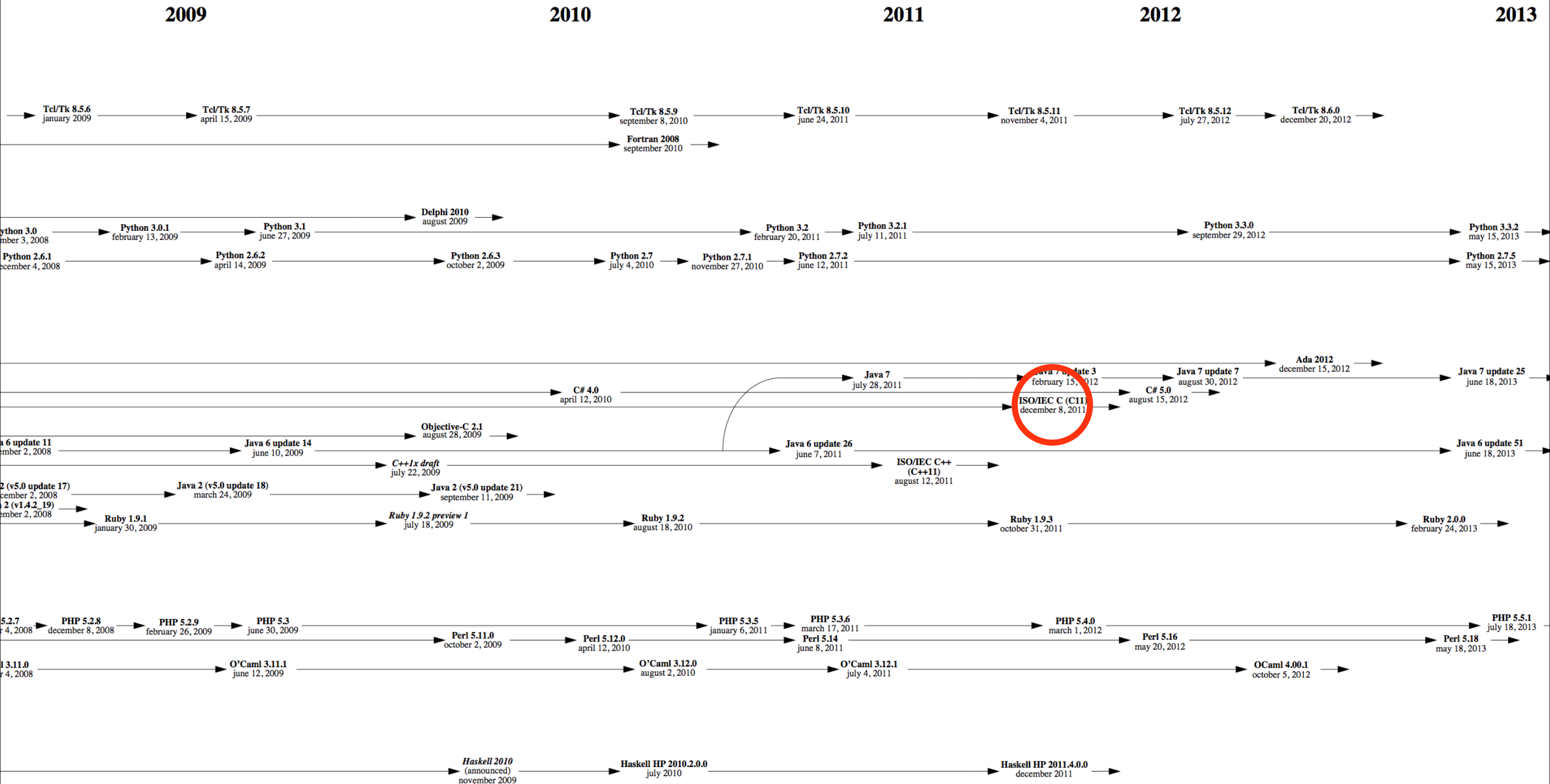
Evolution



Evolution



Evolution



Bücher

- „The C Programming Language“ (2nd Edition)
 - Brian W. Kernighan
 - Dennis M. Ritchie

C11

Standard

Warum Standardisierung?

- C entwickelt sich schnell
 - Entspricht nicht mehr dem von Kernighan und Ritchie beschriebenen Spezifikation
 - Viele Erweiterungen, fehlende Vereinheitlichung
- Ziel ist eine Normierung der Sprache

C Standard

- C wurde mehrfach standardisiert
 - C89/C90, C95, C99, C11
- 1983 ANSI (American National Standards Institute) Komitee X3J11 nimmt sich der Ausarbeitung eines Standards an
- 1989 erster Standard „ANSI X3.159-1989 Programming Language C“
- 1990 übernimmt ISO den Standard mit kleinen Änderungen

C Standard

- 1995 ergänzt ISO den Standard C90
- 1999 wird ISO/IEC 9899 verabschiedet (C99), in diesen Standard fließen aus C++ bekannte Erweiterungen zurück in C
- Seitdem Arbeit an C1X durch Komitee WG14, heute bekannt als C11

Regeln für Änderungen

- Vollständige Kompatibilität zu vorherigem Standard
- Bestehender Code hat hohen Stellenwert
- Modifikationen sollen Geist der Sprache nicht verändern
- Das Prinzip des Vertrauens (z.B. Programmierer muss sicherstellen, dass der Speicher das Ergebnis einer Operation aufnehmen kann), wird durch optionale Erweiterungen ergänzt
 - z.B. Bounds-Checking (Annex K, folgt später)

Features

Überblick

- Thread Unterstützung
- Atomare Operationen
- Bounds-Checking
- Static Assertions
- Weitere Neuerungen
- Compiler-Unterstützung

Multithreading

- `<threads.h>` Bestandteil von C seit C11
- Zuvor Rückgriff auf
 - POSIX Threads (im UNIX Umfeld)
 - Windows Threads
- Anlehnung an POSIX Threads
 - Bereits gut dokumentiert
 - Anwender nicht verwirren

Multithreading

- Problem vor C11
 - C Spezifikation sollte eine abstrakte Maschine referenzieren, die eine Generalisierung der benutzten Maschine darstellt
 - Bezieht sich weder auf einen speziellen Compiler, noch auf ein Betriebssystem, noch auf eine CPU
 - Programmiert man entsprechend der Spezifikationen, sollte ein Programm zu jeder Zeit auf jedem System ohne Modifikationen kompilierbar und ausführbar sein

Multithreading

- Problem vor C11
 - Die abstrakte Maschine, die C referenziert, ist vor C11 singlethreaded
 - Es gibt folglich keine Spezifikation für Multithreading und das Programm ist nicht auf jeder Maschine ohne Modifikationen ausführbar
 - Man unterscheidet bei Multithreading sogar zwischen Betriebssystemen
- C11 löst mit einer neuen Spezifikation dieses Problem

Multithreading

- **`int thrd_create(thrd_t *thr, thrd_start_t func, void *arg);`**
 - `thrd_create` erzeugt einen neuen Thread
 - Führt `func(arg)` aus
 - Bei Erfolg wird `thr` der Identifier des neu erzeugten Threads

Multithreading

- **_Noreturn void thrd_exit(int res);**
 - `thrd_exit` beendet einen Thread
 - Gibt das Resultat an `res`
 - Das Programm terminiert nach beenden des letzten Threads normal

Multithreading

- **`int thrd_join(thrd_t thr, int *res);`**
 - `thrd_join` führt den aktuellen Thread mit dem Thread `thr` durch blockieren zusammen, bis der Thread terminiert ist
 - Wenn `res` kein Nullpointer ist, wird in `res` das Resultat des Threads gespeichert

Atomare Operationen

- Beispiel **vor** C11 Spezifikation
 - `int a; int b;`

Thread 1	Thread 2
<code>a = 1;</code>	<code>printf(„a = %d“, a);</code>
<code>b = 2;</code>	<code>printf(„b = %d, b);</code>

- **Vor** C11 ist das Verhalten undefiniert, da es keine Spezifikation für Threads gibt

Atomare Operationen

- Problemstellung: Mehrere Threads sollen die selben Variablen nutzen
- Operationen sollen serialisiert werden
- Ausführungsreihenfolge kann festgelegt werden
- Vor C11 bereits möglich durch Rückgriff auf externe Bibliotheken
 - z.B. GLib

Atomare Operationen

- Mutex
 - `mtx_lock(&mutex);`
 - `a=1;`
 - `mtx_unlock(&mutex);`
- Für einfache Operationen zu komplex
- Es gibt Prozessoren die solche Operationen als einen Schritt durchführen
 - Diese Operationen sind teurer, da Speicherzugriff für andere Prozesse kurzzeitig blockiert werden muss

Atomare Operationen

- C11 abstrahiert diese Prozessorbefehle
- `<stdatomic.h>` stellt Typen, Makros und Funktionen für atomare Operationen bereit
- Atomare Operationen werden häufig im Linux-Kernel verwendet, da sie performanter als Lockingmechanismen sind

Atomare Operationen

- **`void atomic_init(volatile A *obj, C value);`**
 - `atomic_init` initialisiert ein atomares Objekt des Zeigers `obj` mit dem Wert `value`
 - Beispiel
 - `atomic_int a;`
 - `atomic_init(&a, 1);`

Atomare Operationen

- `void atomic_store(volatile A *object, C desired);`
- `void atomic_store_explicit(volatile A *object, C desired, memory_order order);`
- `atomic_store` ersetzt den Wert von `object` mit dem Wert aus `desired`
- `order` gibt das Verhalten beim Ersetzen an

Atomare Operationen

- **C `atomic_load(volatile A *object);`**
- **C `atomic_load_explicit(volatile A *object, memory_order order);`**
 - `atomic_load` lädt den Wert eines Objektes und gibt ihn zurück
 - `order` gibt an wie der Wert geladen werden soll

Atomare Operationen

- Beispiel mit C11 Spezifikation
 - `atomic_int a; atomic_int b;`

Thread 1	Thread 2
<code>atomic_store(&a, 1);</code>	<code>atomic_load(&b);</code>
<code>atomic_store(&b, 2);</code>	<code>atomic_load(&a);</code>

- Seit der C11 Spezifikation ist das Verhalten definiert

Atomare Operationen

- Es gibt 3 mögliche Verhalten, **Thread2** lädt
 - (1) **b=0** und **a=0**, wenn dieser **vor Thread1** ausgeführt wird
 - (2) **b=2** und **a=1**, wenn dieser **nach Thread1** ausgeführt wird
 - (3) **b=0** und **a=1**, wenn dieser ausgeführt wird **nachdem Thread1 a** gespeichert hat und **bevor b** gespeichert wurde

Atomare Operationen

- Was nicht passieren kann ist
 - **b=2** und **a=0**, weil die Standardeinstellungen für atomares Laden/Speichern sequentielle Konsistenz vorschreibt
 - Das bedeutet, dass sie in der Reihenfolge geschehen müssen wie sie in den Threads beschrieben sind

Atomare Operationen

- Nachteil
 - Dieses Verhalten ist auf einer CPU sehr teuer
- Lösung
 - Wenn der Algorithmus keine Ordnung benötigt, also beispielsweise auch den Zustand **b=2** und **a=0** akzeptiert, kann für Laden/Speichern eine bestimmte Ordnung angegeben werden

Atomare Operationen

- Beispiel
 - `atomic_int a; atomic_int b;`

Thread 1	Thread 2
<code>atomic_store(&a, 1, memory_order_relaxed)</code>	<code>atomic_load(&b, memory_order_relaxed)</code>
<code>atomic_store(&b, 2, memory_order_relaxed)</code>	<code>atomic_load(&a, memory_order_relaxed)</code>

- Dieses Verhalten ist auf modernen CPUs wesentlich schneller

Atomare Operationen

- Zusammenfassend standardisiert C11 Mutexe
- C11 bietet in seiner Spezifikation Möglichkeiten an, wie beim Laden/Speichern vorzugehen ist
- Resultat
 - Es lassen sich performante Routinen entlang der Spezifikationen schreiben
 - Der Code lässt sich jetzt und in Zukunft auf allen Systemen ohne Modifikationen ausführen

Bounds-Checking

- Annex K des ISO/IEC 9899:2011 Standards
- Sicherheitsrelevante Erweiterungen
- Sicherstellung von genügend Speicher für das Ergebnis einer Operation lag zuvor beim Entwickler
 - Folge: z.B. Buffer-Overflows
- Seit C11 unterstützende Methoden mit Suffix `_s`, um vor Sicherheitslücken zu schützen

Bounds-Checking

- Beispiel `strcat`
 - **`errno_t strcat_s(char * restrict s1, rsize_t s1max, const char * restrict s2);`**
 - Erwartet einen zusätzlichen Parameter für die Puffergröße
 - `strcat_s` kopiert nicht mehr als `s1max` Bytes zu `s1`

Bounds-Checking

- Beispiel strcpy
 - **errno_t strcpy_s(char * restrict s1, rsize_t s1max, const char * restrict s2);**
 - `strcpy_s` erfordert, dass `s1max` größer ist als die Länge von `s2`

Bounds-Checking

- Beispiel `gets`
 - **`char *gets_s(char * restrict buffer, size_t nch);`**
 - `gets_s` liest höchstens `nch` Zeichen von der Standardeingabe
 - Zuvor war `gets` eine häufig genutzte Sicherheitslücke
 - In C99 wurde es bereits deprecated
 - In C11 komplett entfernt

Bounds-Checking

- Methoden angelehnt an Visual C
 - Dort gab es Methoden mit Suffix `_s` schon vor C11
 - Implementierung in C11 ähnlich, aber nicht identisch

Static Assertions

- Ausdruck, der zur Übersetzungszeit ausgewertet wird und bei Auswertung gleich **0** einen String/Fehler ausgibt
- Hilfreich zum Debuggen/Fehleranalyse
- In Java bereits lange vorhanden

Static Assertions

- Zur Zeit der C89 Spezifikation wurden Tests vom Präprozessor ausgeführt
 - Ausgabe der Fehler durch Präprozessoranweisung `#error`
- Tests durch Präprozessor sind sehr limitiert
 - `sizeof` wird beispielsweise nicht korrekt ausgewertet
 - Das liegt daran, dass diese Methoden erst konvertiert werden wenn der Präprozessor ausgeführt wurde

Static Assertions

- C89 Beispiele
 - `#if __STDC__ != 1`
 - `#error "Fehlerbeschreibung"`
 - `#endif`

 - `#if sizeof(long) < 8`
 - `#error "Fehlerbeschreibung"`
 - `#endif`

Static Assertions

- Benötigt wird eine Anweisung, die diese Funktion zur Übersetzungszeit übernimmt (nachdem der Präprozessor ausgeführt wurde)
 - C11 stellt `_Static_assert` bereit

Static Assertions

- `_Static_assert` besteht aus
 - Einem Ausdruck, der zur Übersetzungszeit ausgewertet wird
 - Einem String, der ausgegeben wird, wenn die Anweisung zu `0` ausgewertet wird
- **`_Static_assert(constant-expression, string-literal);`**

Static Assertions

- Beispiel
 - `_Static_assert(sizeof(size_t) >= 8, "Fehlerbeschreibung");`

Weitere Neuerungen

- Exklusiver Dateizugriff
- Generic-Selections
- Unicode-Unterstützung
- Speicherzugriffe und Speicherausrichtung
- ...

Compiler-Unterstützung

- Nicht alle Compiler unterstützen die C11 Spezifikation vollständig
- Gilt ebenfalls, auch heute noch, für C99

Compiler-Unterstützung

	C99	C11
LLVM/ Clang	Nahezu alle Features (keine C99 Floating Point Pragmas)	Teilweise
Pelles C	Vollständige Unterstützung	Vollständige Unterstützung
IBM XL	Vollständige Unterstützung	Vollständige Unterstützung
GCC	Nahezu alle Features (1 Feature falsch, 6 nicht vorhanden)	Teilweise

Quellen

Quellen

- <http://www.open-std.org/jtc1/sc22/wg14/www/docs/n1570.pdf>
- <http://stackoverflow.com/questions/6319146/c11-introduced-a-standardized-memory-model-what-does-it-mean-and-how-is-it-g>
- <http://bartoszmilewski.com/2008/12/01/c-atomics-and-memory-ordering/>
- <http://blog.smartbear.com/codereviewer/c11-a-new-c-standard-aiming-at-safer-programming/>

Quellen

- <https://www.securecoding.cert.org/confluence/display/seccode/STR07-C.+Use+the+bounds-checking+interfaces+for+remediation+of+existing+string+manipulation+code>
- <http://www.ibm.com/developerworks/rational/library/support-iso-c11/>
- <http://www.smorgasbordet.com/pellesc/>
- <http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=an&subtype=ca&supplier=897&appname=IBMLinkRedirect&letternum=ENUS202-161>

Quellen

- <http://gcc.gnu.org/c99status.html>
- <http://gcc.gnu.org/wiki/C11Status>