# Abstrakte Interpretation

12./14. Jan. 2005

J.Burghardt, FIRST

#### Übersicht

- Motivation
- Collecting Semantics
- Vom Programm zum Mengengleichungssystem
- Der Fixpunktsatz
- Mit dem Fixpunktsatz zur Lösung

# Übersicht

- Abstrakte Interpretation
- Korrektheit
- Überführbarkeit
- Terminierung
- Aussagekraft

# Übersicht

- Statische Analyse mit PolySpace
- Ergänzungen
- Literatur

#### **Motivation**

Model-Checking-Techniken sind nur anwendbar, wenn das zu verifizierende Programm nur endlich viele Zustände hat.

Sobald z.B. mehrere int-Variablen auftreten, wird der Zustandsraum unendlich bzw. jedenfalls zu groß für Model-Checking.

Wir wollen im Folgenden einen Ansatz entwickeln, der für diese Art von Programmen besser geeignet ist.

# "Collecting Semantics"

Zu einem gegebenen imperativen Programm suchen wir für jeden Punkt im Programmablauf die Menge der dort möglichen Variablenwerte.

Beim Testen würde man das Programm nacheinander mit verschiedenen einzelnen Eingabewerten ablaufen lassen und die auftretenden Variablenwerte beobachten.

Wir wollen stattdessen versuchen, durch Analyse der Programmstruktur die Mengen der auftretenden Werte "auf einen Schlag" zu berechnen.

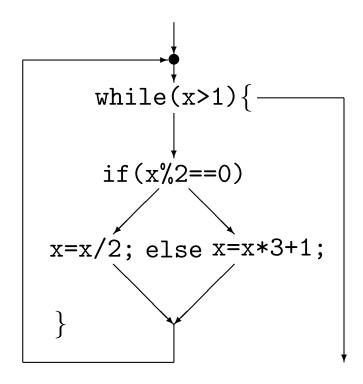
Wir rechnen also nicht mit einzelnen Werten, sondern gleich mit (möglicherweise auch unendlichen) Mengen von Werten.

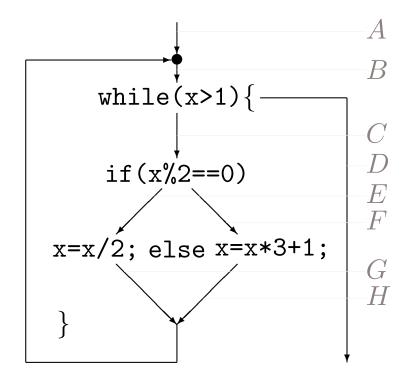
# "Collecting Semantics": Anwendungen

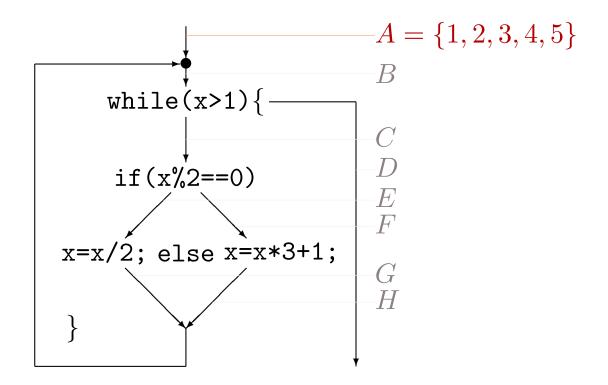
Wenn wir für jede Stelle eines Programms die Menge der dort möglichen Variablenwerte berechnen könnten, wüßten wir alles über die lokalen Eigenschaften des Programms.

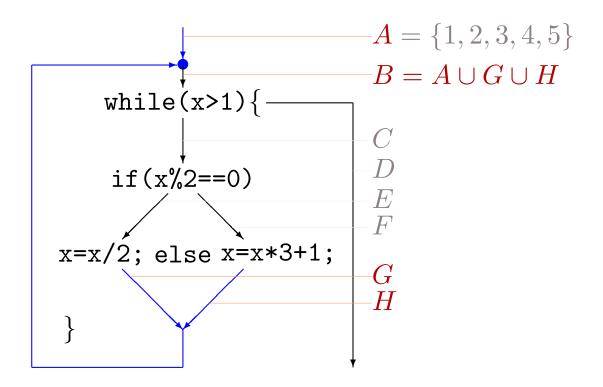
Wir könnten Fragen beantworten wie z.B. kann ein hier ein Uberlauf, eine Division durch Null, ein NULL-Pointer auftreten, ist hier der Feldindex in gültigen Grenzen, wird diese Stelle überhaupt erreicht.

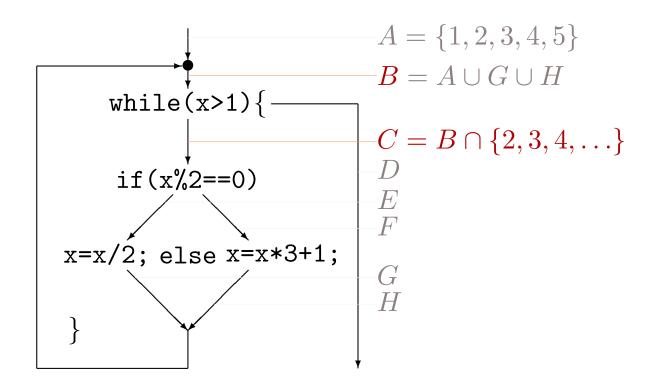
Wir könnten aber nichts über Terminierung und nichts über den Zusammenhang zwischen Ein- und Ausgabe sagen.

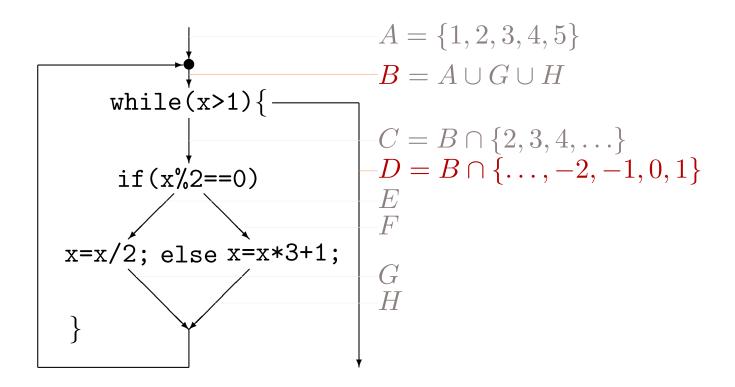


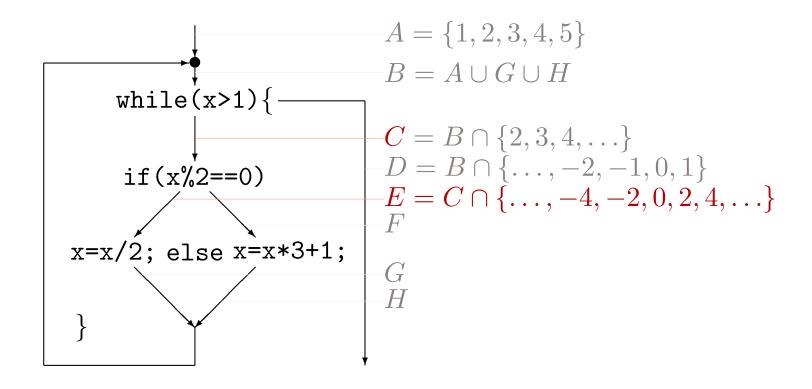


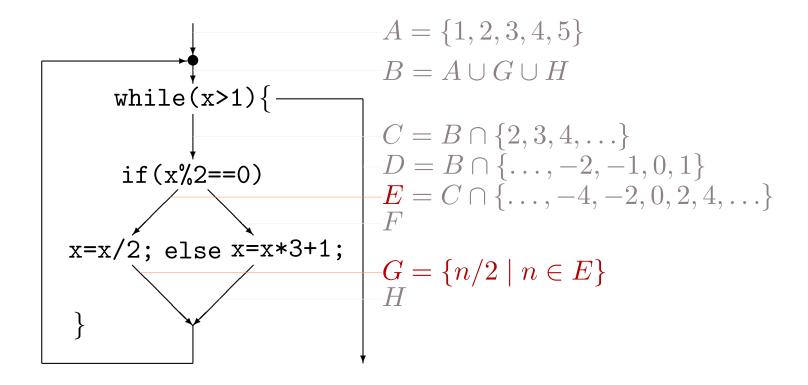


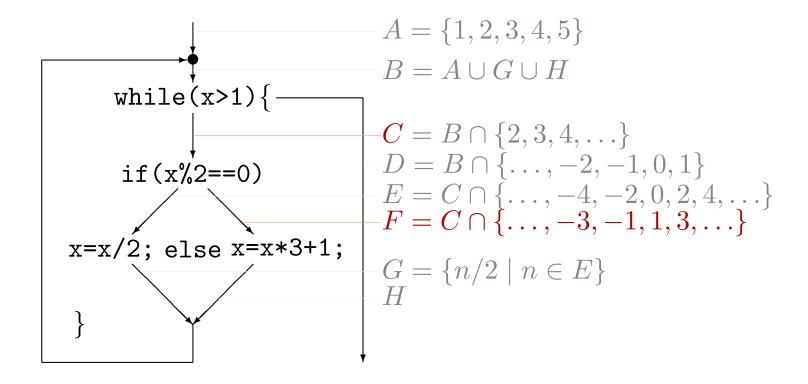


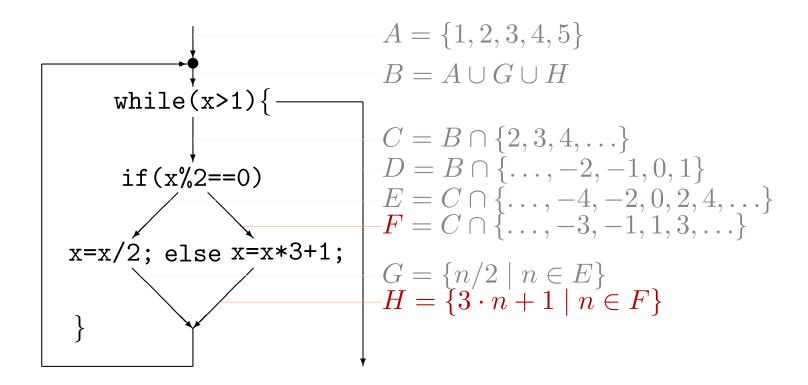


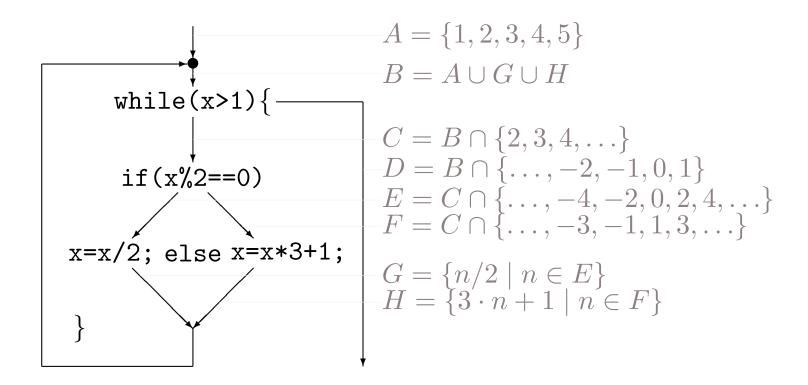






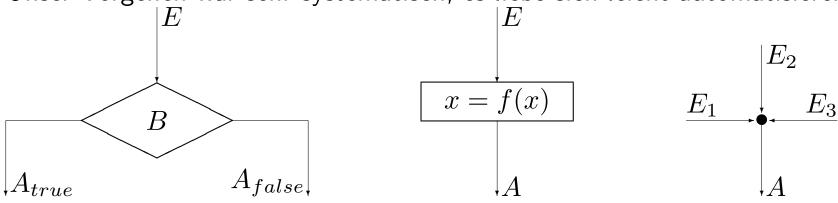






Wir haben aus einem (Spielbeispiel-)Programm ein Mengen-Gleichungssystem abgeleitet.

Unser Vorgehen war sehr systematisch, es ließe sich leicht automatisieren:



$$A_{true} = E \cap \{x \mid B(x)\}$$
  $A = \{f(x) \mid x \in E\}$   $A = E_1 \cup E_2 \cup E_3$   
 $A_{false} = E \cap \{x \mid \neg B(x)\}$ 

## Programme mit mehreren Variablen

In unserem Beispielprogramm tritt nur eine Variable (int x;) auf.

Programme mit mehreren Variablen (z.B. int x, y z;) lasssen sich durch Verwendung von Record-Typen auf eine Variable zurückführen (z.B. struct { int x,y,z } vars;).

Wir werden sie daher hier nicht gesondert behandeln.

Hat ein solches Gleichungssystem zwischen Mengen eine Lösung?

Wie können wir sie finden?

Hat ein solches Gleichungssystem zwischen Mengen eine Lösung?

Wie können wir sie finden?

#### Hat ein solches Gleichungssystem zwischen Mengen eine Lösung?

Das ist zu erwarten: wenn wir das Programm nacheinander mit jedem einzelnen Eingabewert ablaufen lassen und an jedem Programmpunkt A,..., H die beobachteten Werte sammeln (daher "collecting semantics"), sollten wir sie finden.

Wie können wir sie finden?

#### Hat ein solches Gleichungssystem zwischen Mengen eine Lösung?

Das ist zu erwarten: wenn wir das Programm nacheinander mit jedem einzelnen Eingabewert ablaufen lassen und an jedem Programmpunkt A,..., H die beobachteten Werte sammeln (daher "collecting semantics"), sollten wir sie finden.

Das geht aber nur, wenn die Eingabemenge endlich ist.

Wie können wir sie finden?

Hat ein solches Gleichungssystem zwischen Mengen eine Lösung? Das ist zu erwarten: wenn wir das Programm nacheinander mit jedem einzelnen Eingabewert ablaufen lassen und an jedem Programmpunkt A,..., H die beobachteten Werte sammeln (daher "collecting semantics"), sollten wir sie finden.

Das geht aber nur, wenn die Eingabemenge endlich ist.

Wie können wir sie finden?

Hat ein solches Gleichungssystem zwischen Mengen eine Lösung? Das ist zu erwarten: wenn wir das Programm nacheinander mit jedem einzelnen Eingabewert ablaufen lassen und an jedem Programmpunkt A,..., H die beobachteten Werte sammeln (daher "collecting semantics"), sollten wir sie finden.

Das geht aber nur, wenn die Eingabemenge endlich ist.

Wie können wir sie finden?

Hat es eine eindeutige Lösung?

Nein.

Auf der nächsten Folie werden mehrere Lösungen gezeigt.

Wir müssen uns klar werden, welche die "richtige" ist.

Hat ein solches Gleichungssystem zwischen Mengen eine Lösung? Das ist zu erwarten: wenn wir das Programm nacheinander mit jedem einzelnen Eingabewert ablaufen lassen und an jedem Programmpunkt A,..., H die beobachteten Werte sammeln (daher "collecting semantics"), sollten wir sie finden.

Das geht aber nur, wenn die Eingabemenge endlich ist.

Wie können wir sie finden?

Hat es eine eindeutige Lösung? Nein.

Auf der nächsten Folie werden mehrere Lösungen gezeigt. Wir müssen uns klar werden, welche die "richtige" ist.

#### Gleichungssystem

$$A = \{1, 2, 3, 4, 5\}$$

$$B = A \cup G \cup H$$

$$C = B \cap \{2, 3, 4, ...\}$$

$$D = B \cap \{..., -2, -1, 0, 1\}$$

$$E = C \cap \{..., -2, 0, 2, ...\}$$

$$F = C \cap \{..., -1, 1, 3, ...\}$$

$$G = \{n/2 \mid n \in E\}$$

$$H = \{3 \cdot n + 1 \mid n \in F\}$$

$$\begin{array}{ll} \text{Gleichungssystem} & \text{L\"osung 1} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,4,5,8,10,16\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,5,8,10,16\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,8,10,16\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,4,5,8\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16\} \end{array}$$

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                 B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

#### Gleichungssystem

$$A = \{1, 2, 3, 4, 5\}$$
  $A = \{1, 2, 3, 4, 5\}$   $B = A \cup G \cup H$   $B = \{1, 2, 3, 4, 5\}$   $C = B \cap \{2, 3, 4, ...\}$   $C = \{2, 3, 4, 5\}$   $D = B \cap \{..., -2, -1, 0, 1\}$   $D = \{1\}$   $E = C \cap \{..., -2, 0, 2, ...\}$   $E = \{2, 4, 8, 5\}$   $E = \{n/2 \mid n \in E\}$   $E = \{1, 2, 3, 4, 5\}$   $E = \{1, 2, 4, 5\}$   $E = \{$ 

#### Lösung 2

```
A = \{1, 2, 3, 4, 5\}
B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
A = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
A = \{1, 2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
A = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
A = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
A = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
A = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
B = A \cup G \cup H
                                  B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \ldots\}
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                              G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                  B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \ldots\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                 B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                             D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                  B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \ldots\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                              G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                 B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \ldots\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                  B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                              G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                 B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                             D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, ...\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

```
Gleichungssystem
                                  Lösung 2
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
                                 B = \{1, 2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, \dots\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16, 32, 64, 128, 256, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
                             D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 8, 10, 16, 32, 64, 128, 256, ...\}
F = C \cap \{..., -1, 1, 3, ...\}
                              F = \{3, 5\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 4, 5, 8, 16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}
```

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{...,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{...,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{...,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{...,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{...,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{...,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

```
\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}
```

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{...,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{...,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{...,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

```
\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}
```

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

$$\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{\ldots,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{\ldots,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{\ldots,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}$$

```
\begin{array}{lll} \text{Gleichungssystem} & \text{L\"osung 3} \\ A = \{1,2,3,4,5\} & A = \{1,2,3,4,5\} \\ B = A \cup G \cup H & B = \{1,2,3,\ldots\} \\ C = B \cap \{2,3,4,\ldots\} & C = \{2,3,4,\ldots\} \\ D = B \cap \{...,-2,-1,0,1\} & D = \{1\} \\ E = C \cap \{...,-2,0,2,\ldots\} & E = \{2,4,6,\ldots\} \\ F = C \cap \{...,-1,1,3,\ldots\} & F = \{3,5,7,\ldots\} \\ G = \{n/2 \mid n \in E\} & G = \{1,2,3,\ldots\} \\ H = \{3 \cdot n + 1 \mid n \in F\} & H = \{10,16,22,\ldots\} \end{array}
```

### Gleichungssystem

$$A = \{1, 2, 3, 4, 5\}$$

$$B = A \cup G \cup H$$

$$C = B \cap \{2, 3, 4, ...\}$$

$$D = B \cap \{..., -2, -1, 0, 1\}$$

$$E = C \cap \{..., -2, 0, 2, ...\}$$

$$F = C \cap \{..., -1, 1, 3, ...\}$$

$$G = \{n/2 \mid n \in E\}$$

$$H = \{3 \cdot n + 1 \mid n \in F\}$$

```
Gleichungssystem
                                  Lösung 3
A = \{1, 2, 3, 4, 5\}
                                  A = \{1, 2, 3, 4, 5\}
B = A \cup G \cup H
                                  B = \{1, 2, 3, 4, 5, 8, 10, 16\}32, 64, 128, 256, \dots\}
C = B \cap \{2, 3, 4, ...\}
                              C = \{2, 3, 4, 5, 8, 10, 16\}32, 64, 128, 256, \dots\}
D = B \cap \{..., -2, -1, 0, 1\}
                              D = \{1\}
E = C \cap \{..., -2, 0, 2, ...\}
                              E = \{2, 4, 6, 10\} 16\} 32, 64, 128, 256, ...\}
                              F = \{3, 5\}7, \dots\}
F = C \cap \{..., -1, 1, 3, ...\}
G = \{n/2 \mid n \in E\}
                           G = \{1, 2, 3, 5, 8\}16, 32, 64, 128, 256, \ldots\}
H = \{3 \cdot n + 1 \mid n \in F\} H = \{10, 16\}22, ...\}
```

Wie können wir nun die kleinste Lösung finden?

Wie können wir nun die kleinste Lösung finden?

Wir können einen Fixpunktsatz aus der Verbandstheorie verwenden.

Wie können wir nun die kleinste Lösung finden?

Wir können einen Fixpunktsatz aus der Verbandstheorie verwenden.

Dazu formen wir das Gleichungssystem in einen Operator  $\Phi$  um, so daß jeder Fixpunkt von  $\Phi$ , d.h. jedes  $\langle A,\ldots,H\rangle$  mit  $\Phi(\langle A,\ldots,H\rangle)=\langle A,\ldots,H\rangle$  eine Lösung des Gleichungssystems ist.

Wie können wir nun die kleinste Lösung finden?

Wir können einen Fixpunktsatz aus der Verbandstheorie verwenden.

Dazu formen wir das Gleichungssystem in einen Operator  $\Phi$  um, so daß jeder Fixpunkt von  $\Phi$ , d.h. jedes  $\langle A,\ldots,H\rangle$  mit  $\Phi(\langle A,\ldots,H\rangle)=\langle A,\ldots,H\rangle$  eine Lösung des Gleichungssystems ist.

### Zunächst: vom Gleichungssystem zum Operator

#### Gleichungssystem

```
A = \{1, 2, 3, 4, 5\}
B = A \cup G \cup H
C = B \cap \{2, 3, 4, ...\}
D = B \cap \{..., -2, -1, 0, 1\}
E = C \cap \{..., -4, -2, 0, 2, 4, ...\}
F = C \cap \{..., -3, -1, 1, 3, ...\}
G = \{n/2 \mid n \in E\}
H = \{3 \cdot n + 1 \mid n \in F\}
```

```
Operator
```

```
\Phi: \wp(\mathbf{Z})^{8} \longrightarrow \wp(\mathbf{Z})^{8}, 

\Phi(\langle A, B, C, D, E, F, G, H \rangle) = 

\langle \{1, 2, 3, 4, 5\}, \\
A \cup G \cup H, \\
B \cap \{2, 3, 4, ...\}, \\
B \cap \{..., -2, -1, 0, 1\}, \\
C \cap \{..., -4, -2, 0, 2, 4, ...\}, \\
C \cap \{..., -3, -1, 1, 3, ...\}, \\
\{n/2 \mid n \in E\}, \\
\{3 \cdot n + 1 \mid n \in F\}
```

#### Fixpunkt ist Lösung

$$\Phi(\langle A, B, C, D, E, F, G, H \rangle) = \langle A, B, C, D, E, F, G, H \rangle \Leftrightarrow \{1, 2, 3, 4, 5\} \qquad \land \\ A \cup G \cup H \qquad \land \\ B \cap \{2, 3, 4, \ldots\} \qquad \land \\ B \cap \{\ldots, -2, -1, 0, 1\} \qquad \land \\ C \cap \{\ldots, -4, -2, 0, 2, 4, \ldots\} \qquad \land \\ C \cap \{\ldots, -3, -1, 1, 3, \ldots\} \qquad \land \\ \{n/2 \mid n \in E\} \qquad \land \\ \{3 \cdot n + 1 \mid n \in F\}$$

## Zunächst: vom Gleichungssystem zum Operator

#### **Dipercent long iss**ty Literang

```
\Phi: \wp(\mathbf{Z})^8 \longrightarrow \wp(\mathbf{Z})^8,
\Phi(\langle A, B, C, D, E, F, G, H \rangle) = \langle A, B, C, D, E, F, G, H \rangle \Leftrightarrow
A = \{1, 2, 3, 4, 5\} , \land
B = A \cup G \cup H , \land
C = B \cap \{2, 3, 4, \ldots\} , \land
D = B \cap \{\ldots, -2, -1, 0, 1\} , \land
E = C \cap \{\ldots, -4, -2, 0, 2, 4, \ldots\} , \land
F = C \cap \{\ldots, -3, -1, 1, 3, \ldots\} , \land
G = \{n/2 \mid n \in E\} , \land
H = \{3 \cdot n + 1 \mid n \in F\}
```

## Programme mit mehreren Variablen

Treten mehrere Variablen im Programm auf, wird  $\Phi$  entsprechend komplizierter, z.B.  $\Phi: \wp(\mathbf{Z}^3)^8 \longrightarrow \wp(\mathbf{Z}^3)^8$  für 3 Variablen und 8 Programmstellen.

Z.B. 
$$y=x*y-2$$
; führt zu

$$\Phi(\langle A, \dots, H \rangle) = \langle \dots, \{\langle x, x \cdot y - 2 \rangle \mid \langle x, y \rangle \in E\}, \dots \rangle.$$

Z.B. 
$$if(x < y+1)$$
 führt zu

$$\Phi(\langle A, \dots, H \rangle) = \langle \dots, B \cap \{\langle x, y \rangle \mid x, y \in \mathbf{Z} \land x < y + 1\}, \dots \rangle.$$

## **Der Fixpunktsatz**

Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung.

Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich

$$X = \bigsqcup \{ \bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\Phi(\bot))), \ldots \}.$$

## **Der Fixpunktsatz**

Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung.

Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich  $X = \bigsqcup \ \{\bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\bot)), \ldots\}.$ 

Um diesen Satz verstehen und anwenden zu können, benötigen wir einige formale Definitionen.

## **Der Fixpunktsatz**

Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung.

Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich  $X = \bigsqcup \ \{\bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\bot)), \ldots\}.$ 

Um diesen Satz verstehen und anwenden zu können, benötigen wir einige formale Definitionen.

### **Ordnung**

Sei M eine Menge.

Eine Relation  $(\sqsubseteq) \subseteq M \times M$  heißt *Ordnungsrelation* auf M, wenn für alle  $x,y,z \in M$  gilt:

```
x \sqsubseteq x (Reflexivität)

x \sqsubseteq y \land y \sqsubseteq x \Rightarrow x = y (Antisymmetrie)

x \sqsubseteq y \land y \sqsubseteq z \Rightarrow x \sqsubseteq z (Transitivität).
```

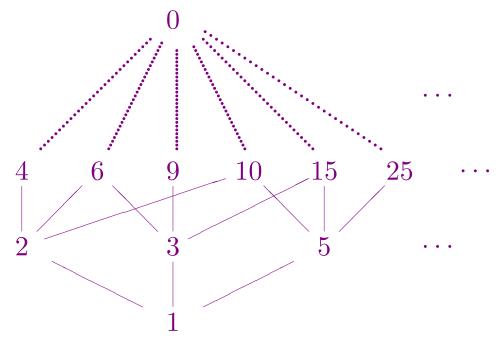
## Beispiele . . .

... aus der Analysis:

 $\boldsymbol{Z}$ ,  $\boldsymbol{Q}$ ,  $\boldsymbol{R}$ , jeweils mit dem üblichen  $(\leq)$ 

## Beispiele . . .

. . . aus der Zahlentheorie:



 $I\!\!N$  mit "x ist Teiler von y"  $(x\mid y)$ 

Das Beispiel zeigt: nicht alle Elemente müssen miteinander vergleichbar sein.

### Beispiele . . .

. . . aus der Mengenlehre:

$$\wp(I\!\!N)$$
 mit  $A_1 \sqsubseteq A_2 \Leftrightarrow A_1 \subseteq A_2$ 

. . . im Hinblick auf unsere Anwendung:

$$\wp(\mathbf{Z})^8 \text{ mit } \langle A_1, \dots, H_1 \rangle \sqsubseteq \langle A_2, \dots, H_2 \rangle \Leftrightarrow A_1 \subseteq A_2 \wedge \dots \wedge H_1 \subseteq H_2$$

Z.B. 
$$\langle \{\}, \{1\}, \{\}, \{\}, \{\}, \{\}, \{3, 5\}, \{\} \rangle \sqsubseteq \langle \{\}, \mathbf{Z}, \{\}, \{\}, \{\}, \{\}, \{3, 4, 5\}, \{\} \rangle$$

### **Schranke**

Ein Element  $x \in M$  heißt *obere Schranke* einer Teilmenge  $S \subseteq M$ , wenn  $s \sqsubseteq x$  für alle  $s \in S$ .

x heißt kleinste obere Schranke von S, wenn  $x \sqsubseteq x'$  für jede obere Schranke x' von S.

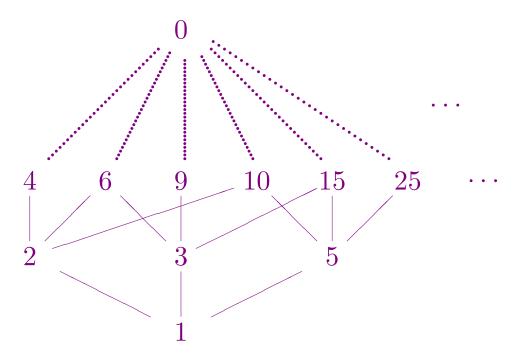
Analog wird "untere Schranke" und "größte untere Schranke" definiert.

 $1.42 \in \mathbf{Q}$  ist eine obere Schranke der Menge  $\{x \in \mathbf{Q} \mid x \cdot x \leq 2\}$ .

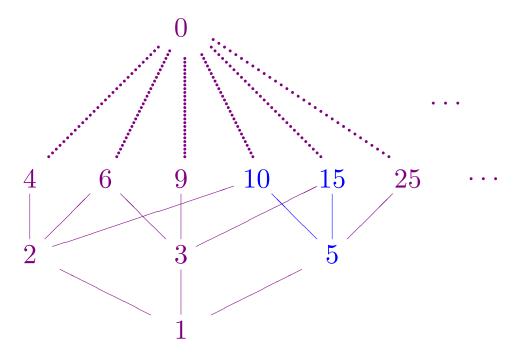
 $1.415 \in \mathbf{Q}$  ist eine kleinere.

Es gibt in *Q* keine kleinste.

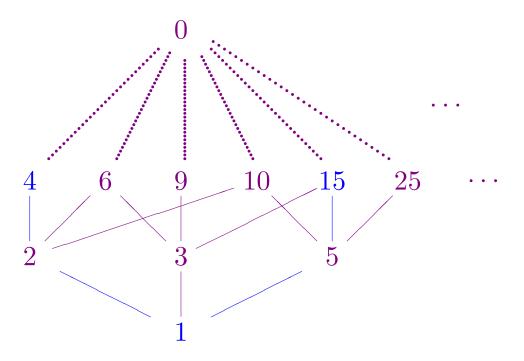
In  $I\!\!R$  gibt es eine kleinste obere Schranke nämlich  $\sqrt{2} \in I\!\!R$ .



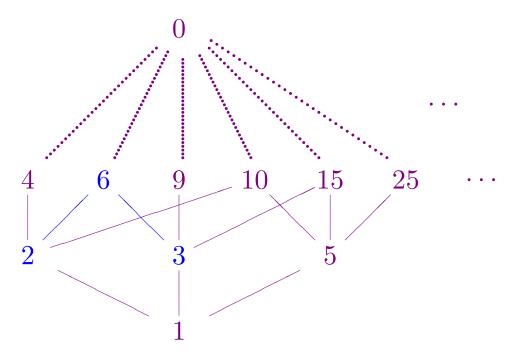
ggT(x,y) bzw. kgV(x,y) ist die größte untere bzw. kleinste obere Schranke von  $\{x,y\}$  bzgl.  $I\!\!N$  mit (|).



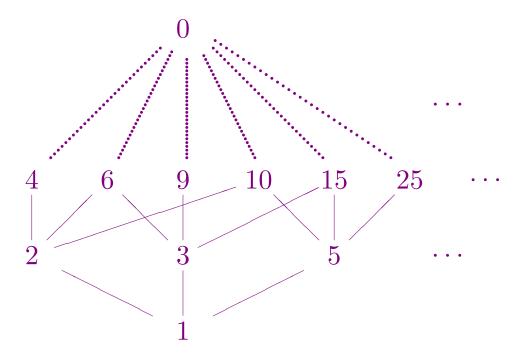
ggT(x,y) bzw. kgV(x,y) ist die größte untere bzw. kleinste obere Schranke von  $\{x,y\}$  bzgl.  $I\!\!N$  mit (|).



ggT(x,y) bzw. kgV(x,y) ist die größte untere bzw. kleinste obere Schranke von  $\{x,y\}$  bzgl.  $I\!\!N$  mit (|).



ggT(x,y) bzw. kgV(x,y) ist die größte untere bzw. kleinste obere Schranke von  $\{x,y\}$  bzgl.  $I\!\!N$  mit (|).



ggT(x,y) bzw. kgV(x,y) ist die größte untere bzw. kleinste obere Schranke von  $\{x,y\}$  bzgl.  $I\!\!N$  mit (|).

 $A_1 \cup \ldots \cup A_n$  bzw.  $A_1 \cap \ldots \cap A_n$  ist kleinste obere bzw. größte untere Schranke von  $\{A_1, \ldots, A_n\}$ .

 $\langle A_1 \cup A_2, \dots, H_1 \cup H_2 \rangle$  bzw.  $\langle A_1 \cap A_2, \dots, H_1 \cap H_2 \rangle$  ist kleinste obere bzw. größte untere Schranke von  $\{\langle A_1, \dots, H_1 \rangle, \langle A_2, \dots, H_2 \rangle\}$ .

#### **Verband**

Die Menge M mit der Ordnungsrelation  $(\sqsubseteq)$  heißt vollständiger Verband, wenn für alle  $S\subseteq M$  stets die kleinste obere und die größte untere Schranke von S existiert.

Wir schreiben  $\bigsqcup S$  bzw.  $\bigsqcup S$  für die kleinste obere bzw. größte untere Schranke von S.

Wir schreiben  $\bot$  bzw.  $\top$  für  $\bigsqcup M$  bzw.  $\square M$ , also für das kleinste bzw. größte Element in M überhaupt.

## Beispiele für vollständige Verbände

```
Z \cup \{-\infty, +\infty\} und R \cup \{-\infty, +\infty\}, jeweils mit dem üblichen (\leq)
\top entspricht +\infty, \perp entspricht -\infty
IN mit "ist Teiler von" (z.B. [1, 3, 5, 7, 9, 11, ...] = 0)
\wp(Z)
\top entspricht Z,
\perp entspricht \{\},
\bigsqcup S entspricht \bigcup S,
\bigcup S entspricht \bigcap S
\wp(\mathbf{Z})^8
\top entspricht \langle \boldsymbol{Z}, \dots, \boldsymbol{Z} \rangle,
\perp entspricht \langle \{\}, \ldots, \{\} \rangle
```

## Keine vollständigen Verbände sind:

Z, Q, R

(haben kein größtes Element)

$$\mathbf{Q} \cup \{-\infty, +\infty\}$$
  
( $\{x \in \mathbf{Q} \mid x \cdot x \leq 2\}$  hat keine kleinste obere Schranke)

## **Monotone Abbildung**

Seien M mit  $(\sqsubseteq)$  sowie M' mit  $(\sqsubseteq')$  vollständige Verbände.

Eine Abbildung  $\Phi: M \longrightarrow M'$  heißt monoton, wenn  $x \sqsubseteq y \Rightarrow \Phi(x) \sqsubseteq' \Phi(y)$  für alle  $x, y \in M$ .

```
f: I\!\!R \cup \{-\infty, +\infty\} \longrightarrow I\!\!R \cup \{-\infty, +\infty\} ist monoton (in unserem Sinne) genau dann, wenn f monoton wachsend im Sinne der Analysis ist. f: I\!\!N \longrightarrow I\!\!N mit f(x) = x \cdot x ist monoton (bzgl. "ist Teiler von"). f(x) = x + 1 ist nicht monoton, denn 6 \sqsubseteq 24, aber f(6) = 7 \not\sqsubseteq 25 = f(24).
```

```
f:\wp(\boldsymbol{Z})\longrightarrow\wp(\boldsymbol{Z}) \text{ mit } f(A)=A\cup\{1,2,3\} \text{ ist monoton (bzgl. } (\subseteq)). f(A)=A\cap\{1,2,3\} \text{ ist monoton.} f(A)=\{10-a\mid a\in A\} \text{ ist monoton,} \text{z.B. ist } \{-1,10\}\sqsubseteq\{-1,10,100\} \text{ und } f(\{-1,10\})=\{11,0\}\sqsubseteq\{11,0,-90\}=f(\{-1,10,100\}). f(A)=\boldsymbol{Z}\setminus A \text{ ist nicht monoton, denn } \{\}\sqsubseteq\boldsymbol{Z}, \text{ aber } f(\{\})=\boldsymbol{Z}\not\sqsubseteq \{\}=f(\boldsymbol{Z}). \Phi:\wp(\boldsymbol{Z})^8\longrightarrow\wp(\boldsymbol{Z})^8 \text{ aus dem "collecting semantics"-Beispiel ist monoton.}
```

## **Stetige Abbildung**

Eine Abbildung  $\Phi: M \longrightarrow M'$  heißt stetig, wenn  $\Phi(\bigsqcup S) = \bigsqcup' \{\Phi(s) \mid s \in S\}$  für jede Teilmenge  $S \subseteq M$ .

```
f: \mathbb{R} \cup \{-\infty, +\infty\} \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\} ist stetig (in unserem Sinn), wenn f(sup(S)) = sup(\{f(x) \mid x \in S\})
```

Unser  $\Phi: \wp(\mathbf{Z})^8 \longrightarrow \wp(\mathbf{Z})^8$  aus dem "collecting semantics" –Beispiel ist stetig.

Allgemeiner: jedes  $\Phi$ , das nur mit Hilfe von  $(\cup)$ ,  $(\cap)$  und  $\{f(x) \mid x \in \ldots\}$  definiert wurde, ist stetig.

Insbesondere ist daher jedes  $\Phi$ , das einem Mengengleichungssystem für ein imperatives Programm entspricht, stetig.

## **Nochmal: Fixpunktsatz**

Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung. Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich  $X = \bigsqcup \{\bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\bot)), \ldots\}.$ 

## **Nochmal: Fixpunktsatz**

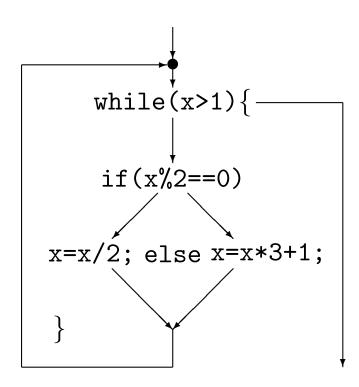
Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung. Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich  $X = \bigsqcup \{\bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\bot)), \ldots\}.$ 

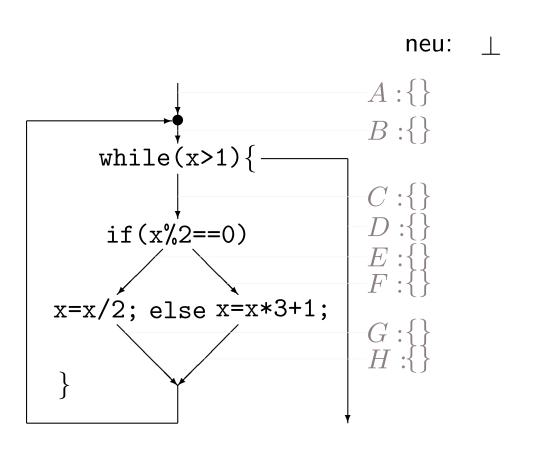
Da unser  $M=\wp(\mathbf{Z})^8$  vollständiger Verband und unser  $\Phi:\wp(\mathbf{Z})^8\longrightarrow\wp(\mathbf{Z})^8$  monotone und stetige Abbildung ist, können wir den Satz anwenden. Der Fixpunkt X hat die Form  $\langle A,\ldots,H\rangle$ ; er ist eine Lösung des ursprünglichen Mengen-Gleichungssystems.

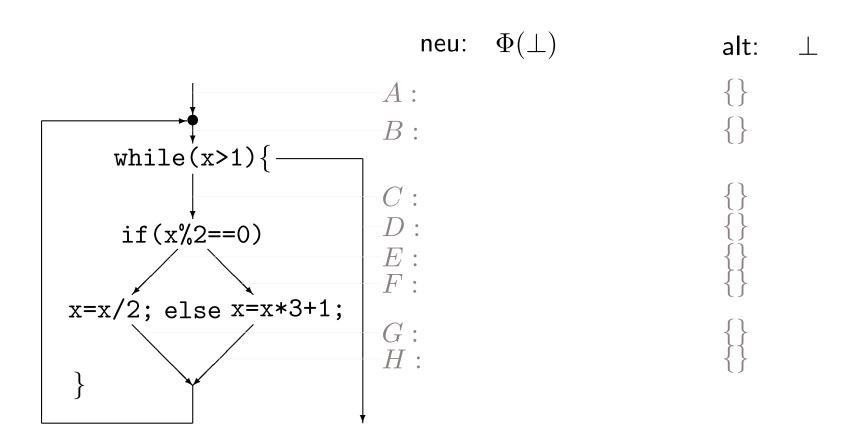
## **Nochmal: Fixpunktsatz**

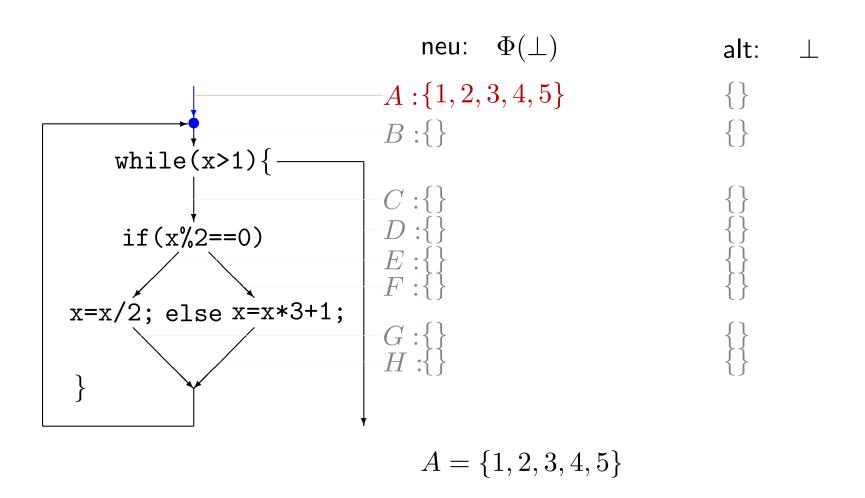
Sei M mit  $(\sqsubseteq)$  ein vollständiger Verband und  $\Phi: M \longrightarrow M$  eine monotone und stetige Abbildung. Dann hat  $\Phi$  einen kleinsten Fixpunkt  $X \in M$ , nämlich  $X = \bigsqcup \{\bot, \ \Phi(\bot), \ \Phi(\Phi(\bot)), \ \Phi(\Phi(\bot)), \ldots\}.$ 

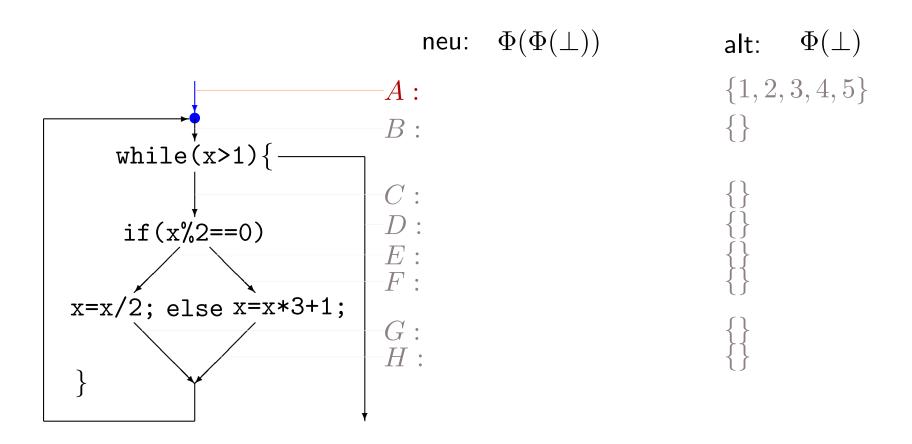
Da unser  $M=\wp(\mathbf{Z})^8$  vollständiger Verband und unser  $\Phi:\wp(\mathbf{Z})^8\longrightarrow\wp(\mathbf{Z})^8$  monotone und stetige Abbildung ist, können wir den Satz anwenden. Der Fixpunkt X hat die Form  $\langle A,\ldots,H\rangle$ ; er ist eine Lösung des ursprünglichen Mengen-Gleichungssystems.

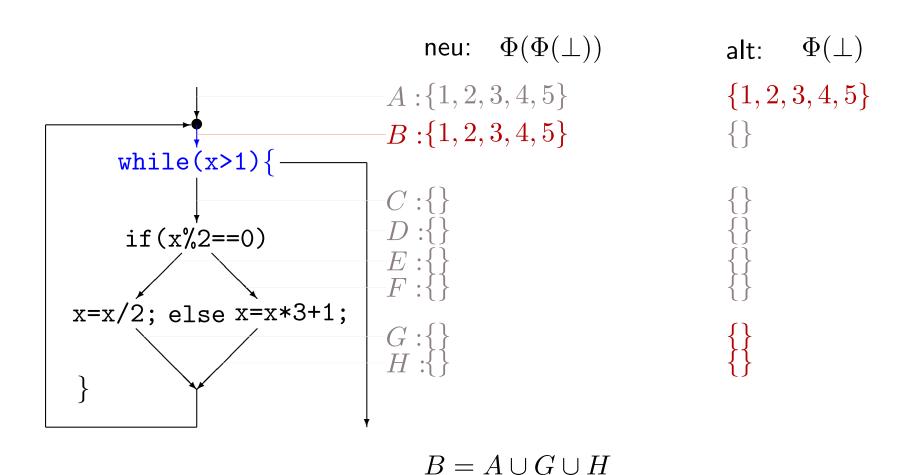


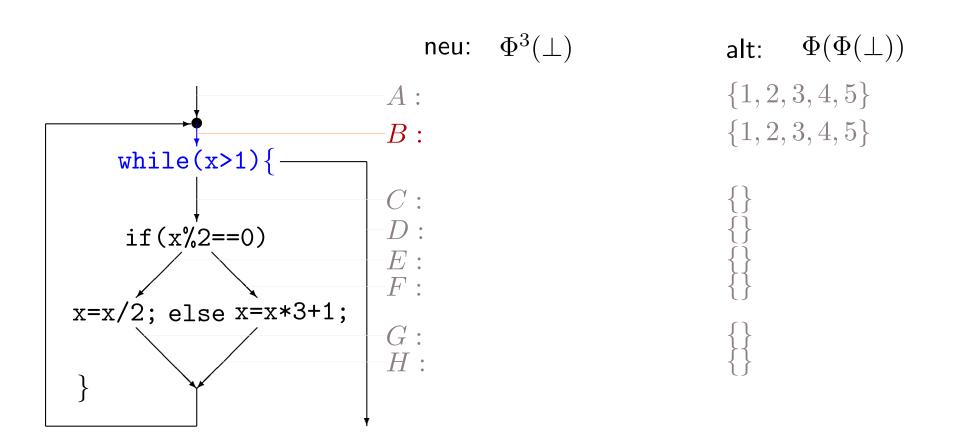


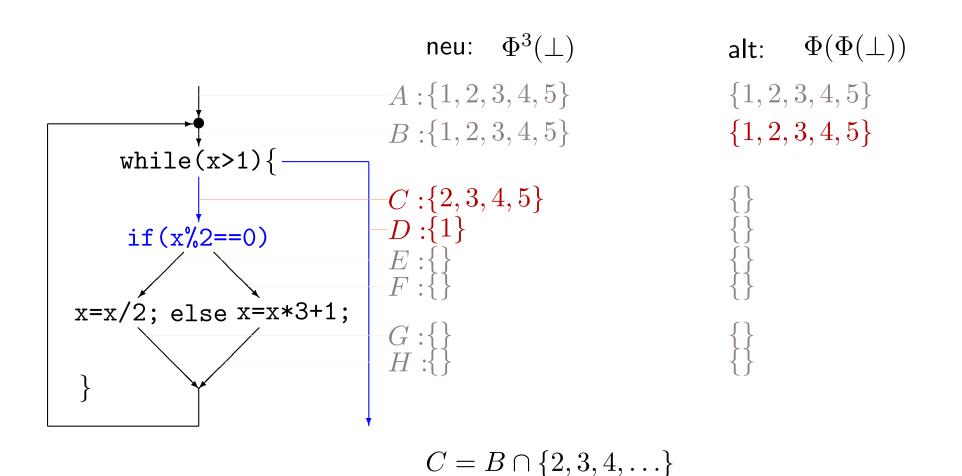


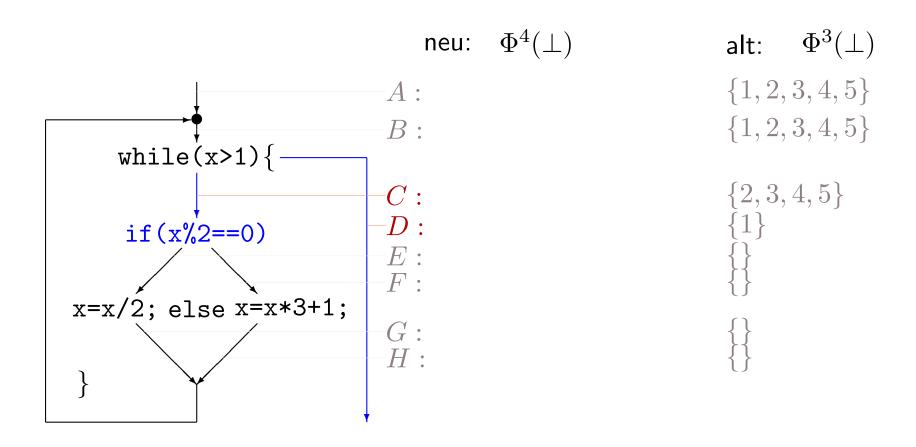


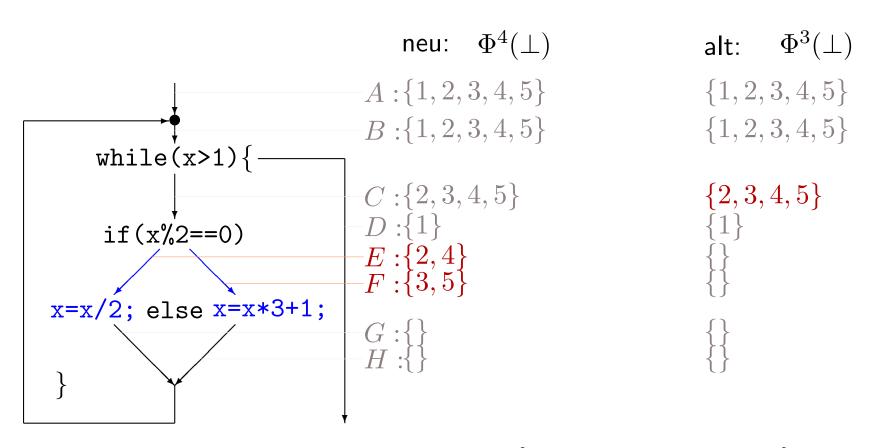






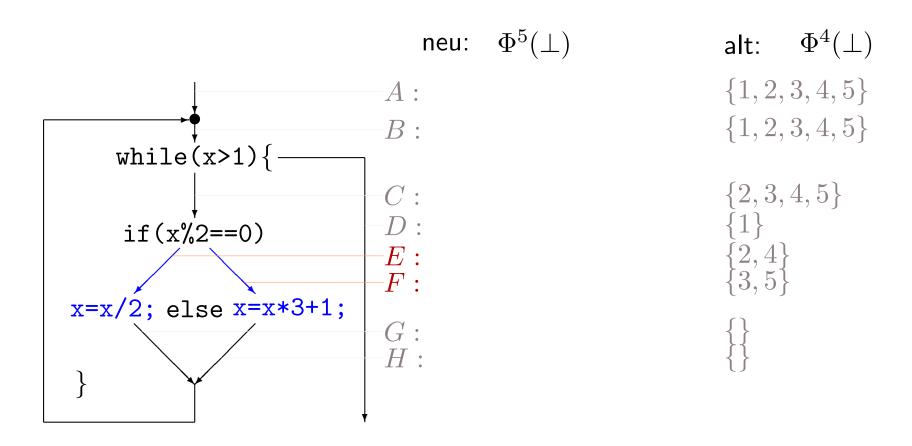


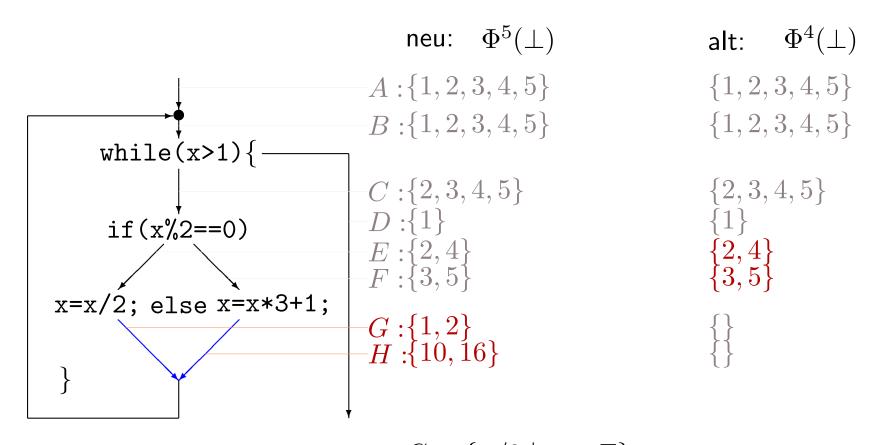




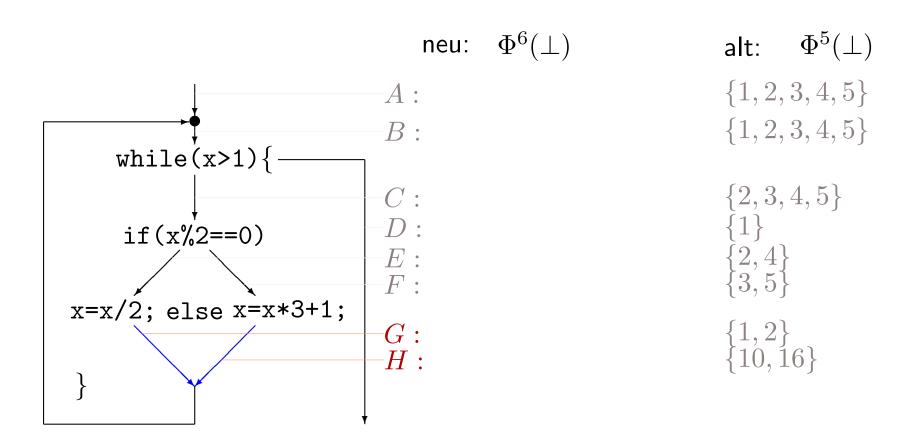
$$E = C \cap \{\dots, -4, -2, 0, 2, 4, \dots\}$$

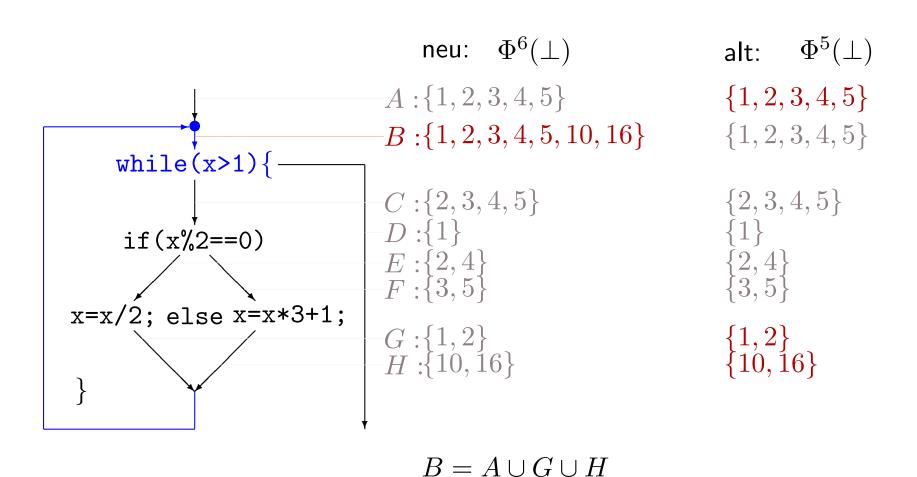
$$F = C \cap \{\dots, -3, -1, 1, 3, \dots\}$$

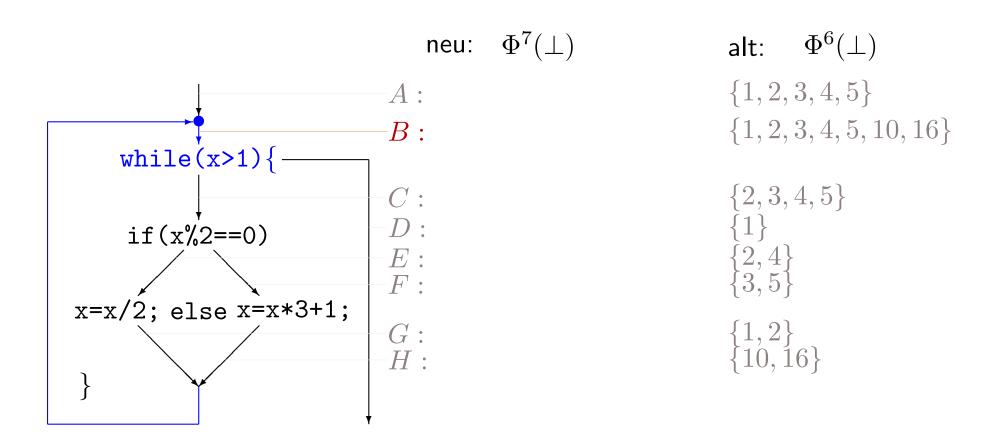


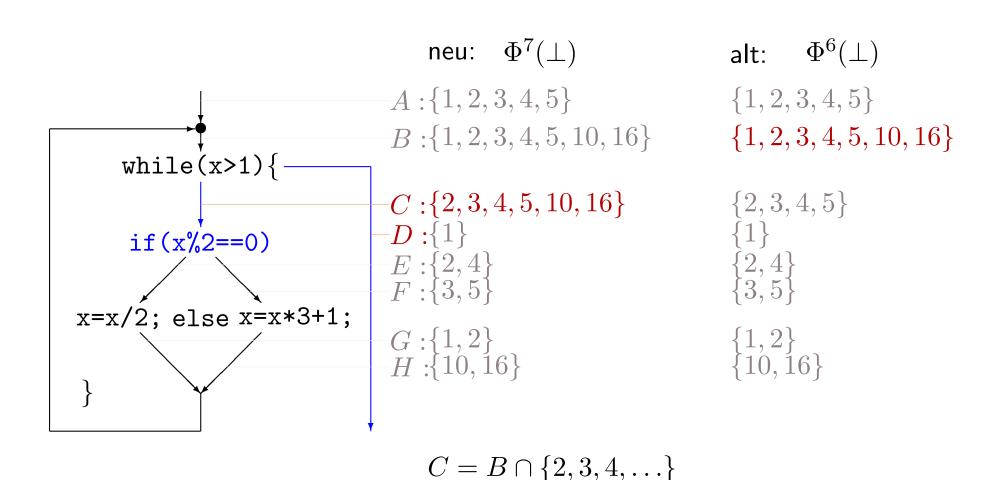


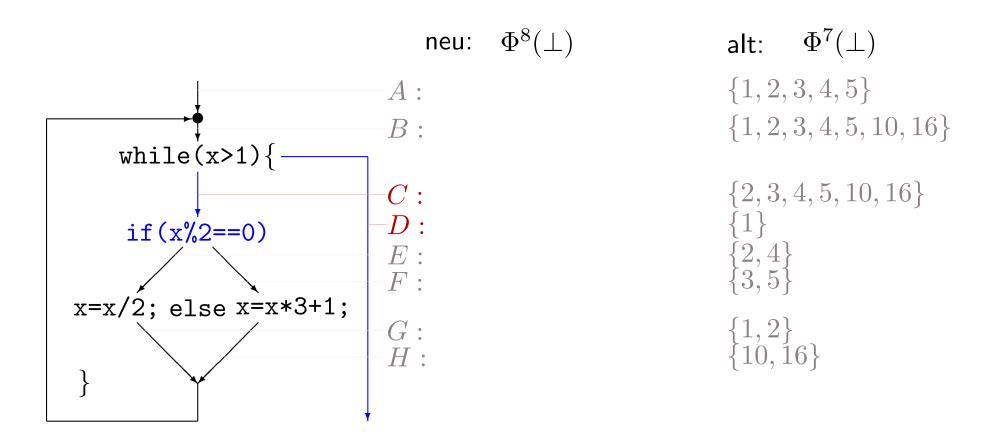
$$G = \{n/2 \mid n \in E\}$$
 
$$H = \{3 \cdot n + 1 \mid n \in F\}$$
 12./14. Jan. 2005 Abstrakte Interpretation

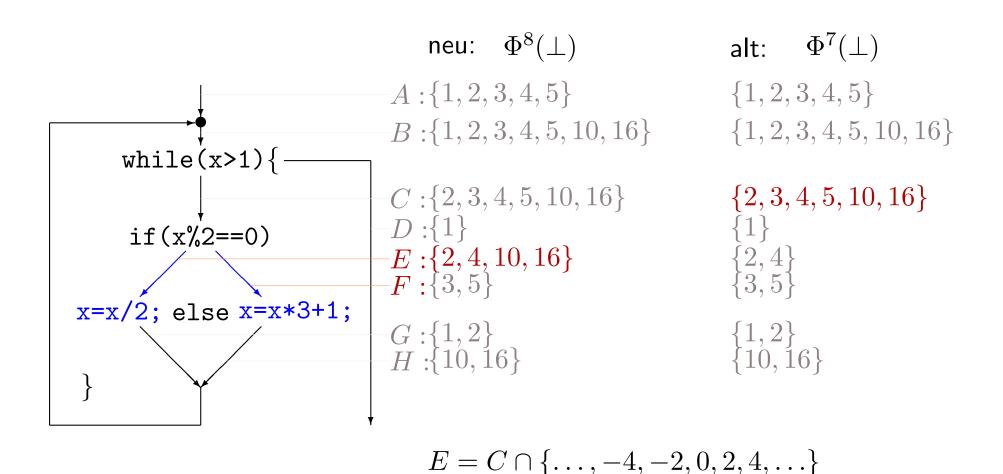


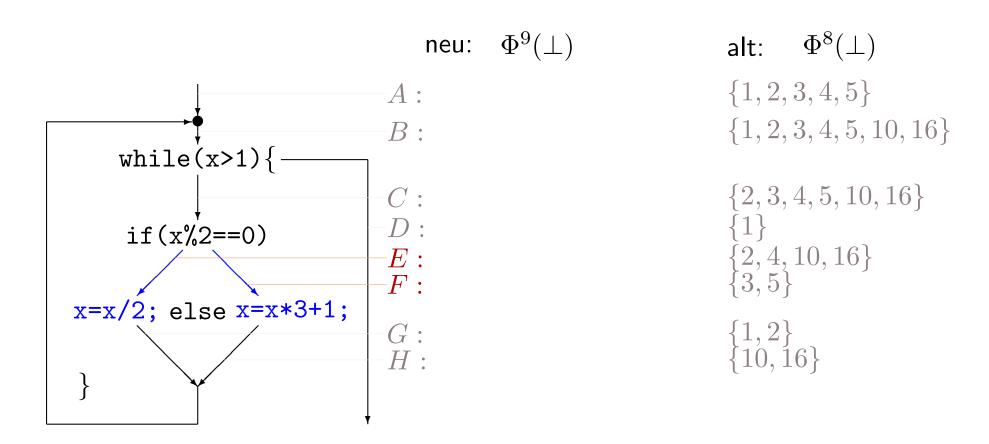


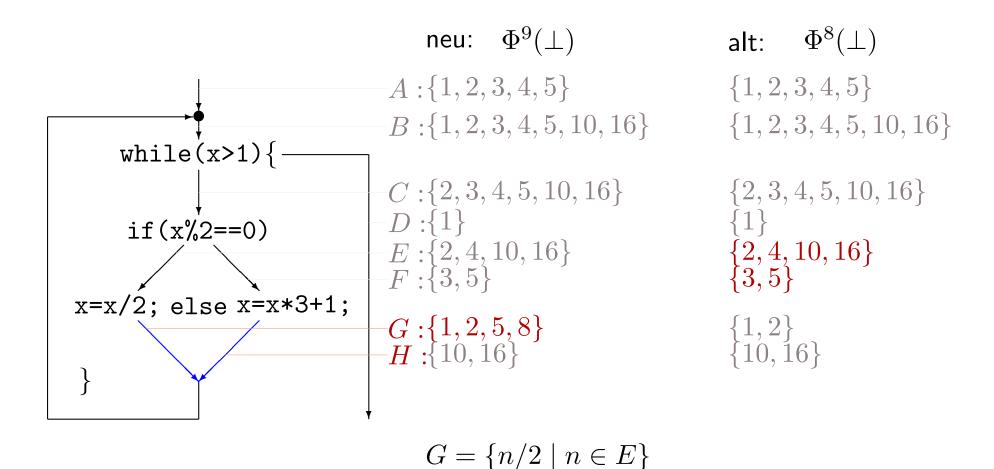


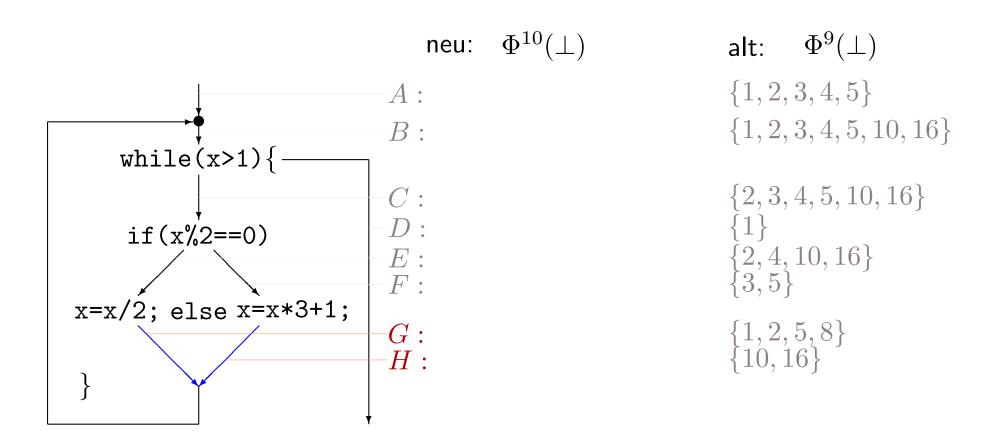


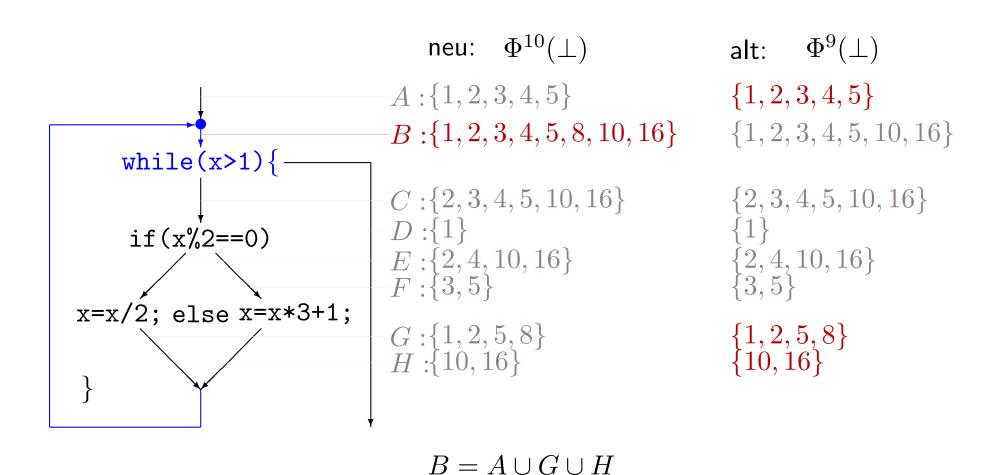


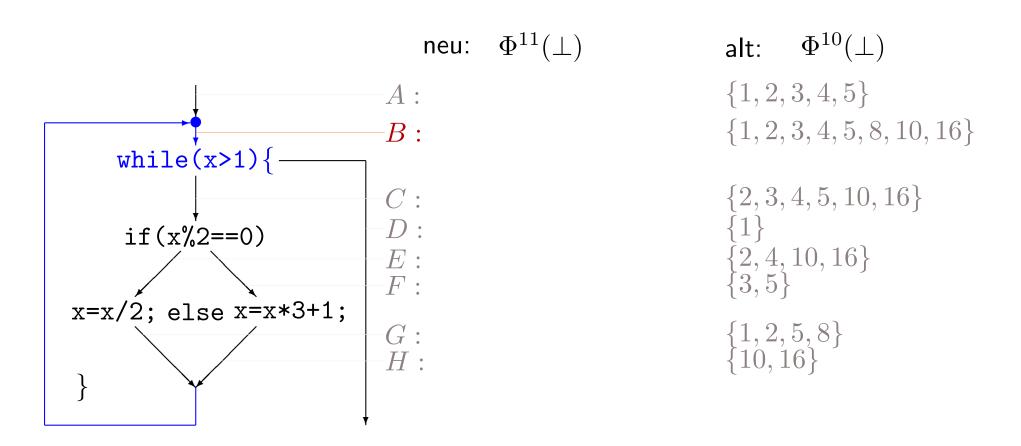


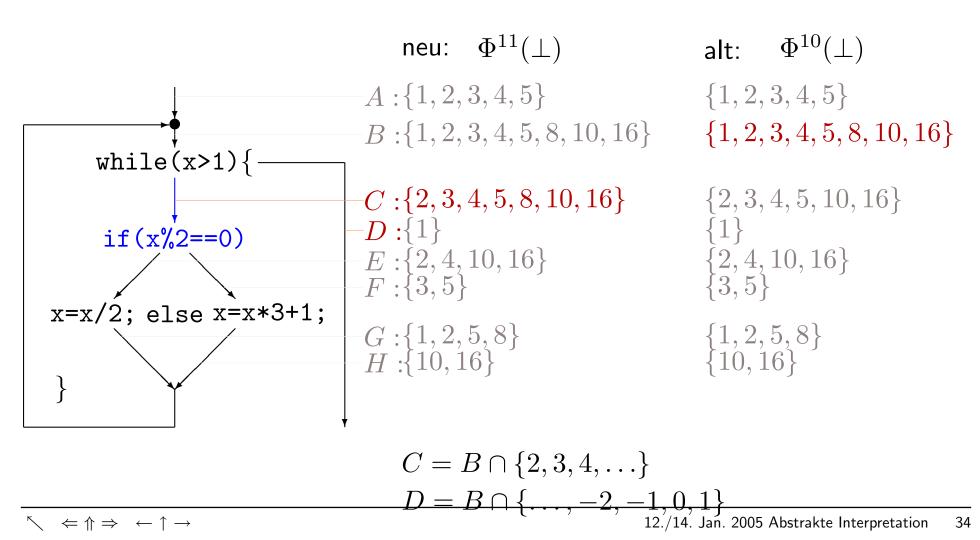


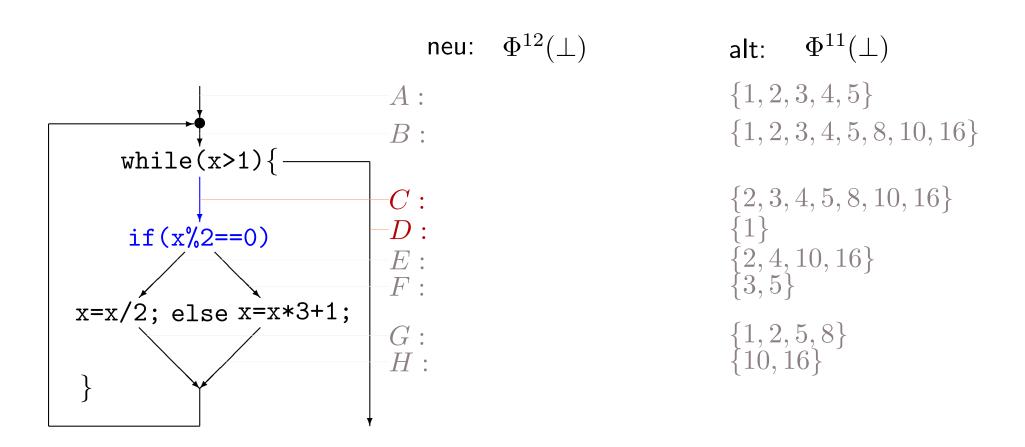


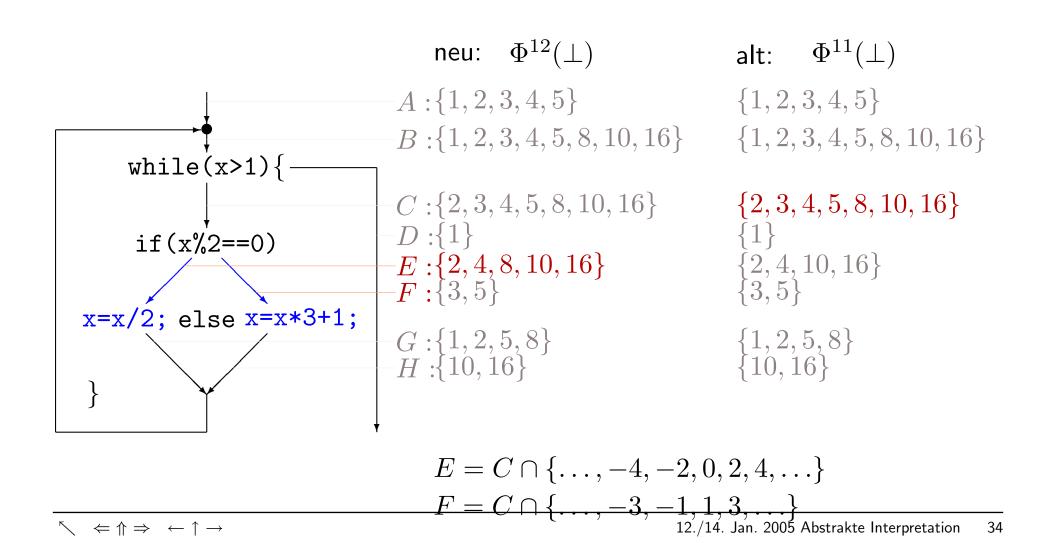


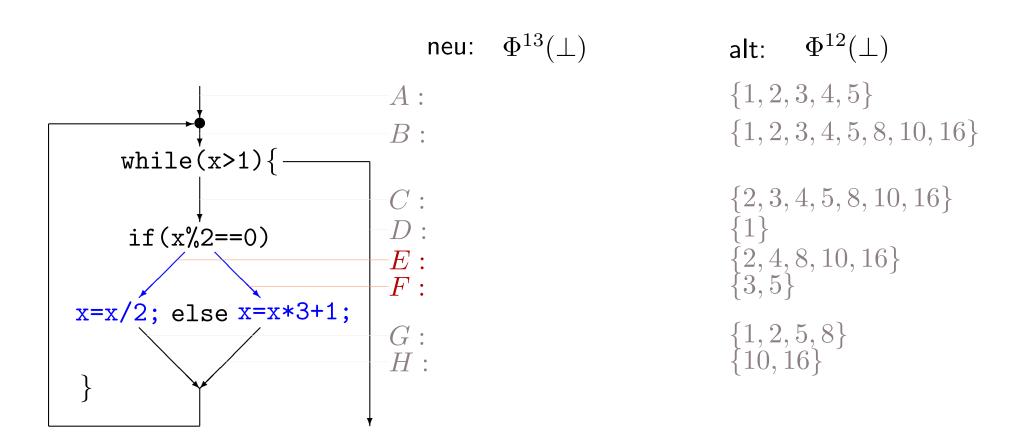


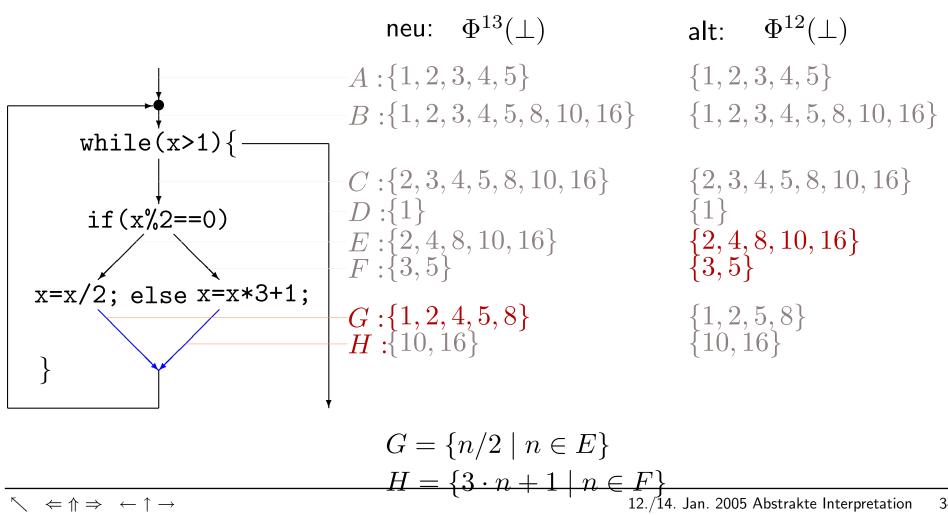


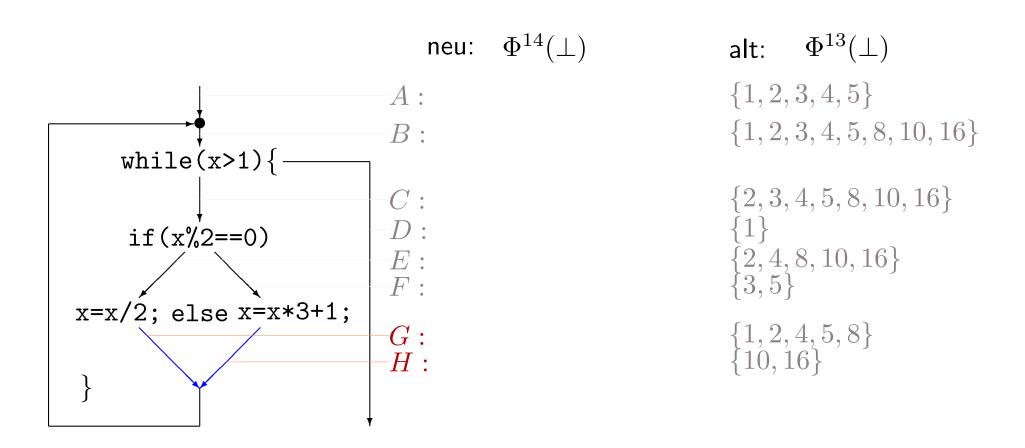


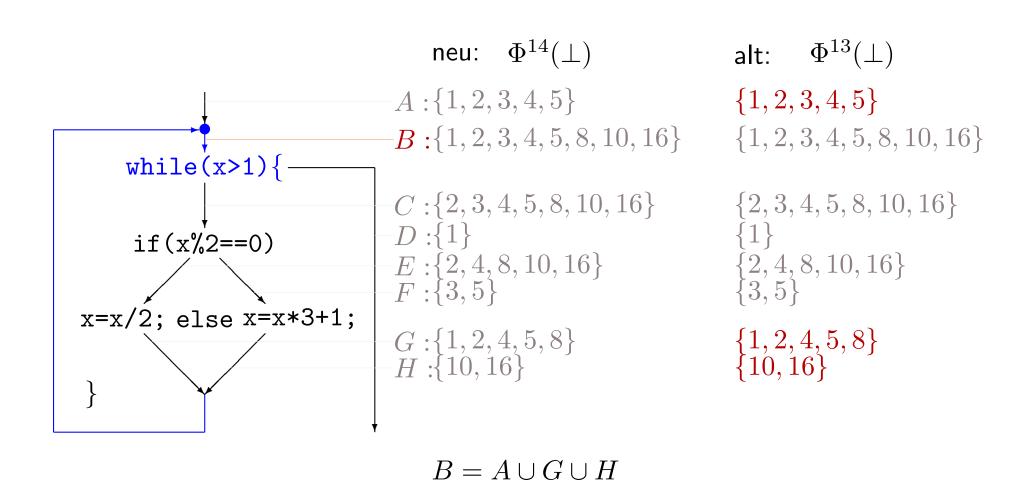


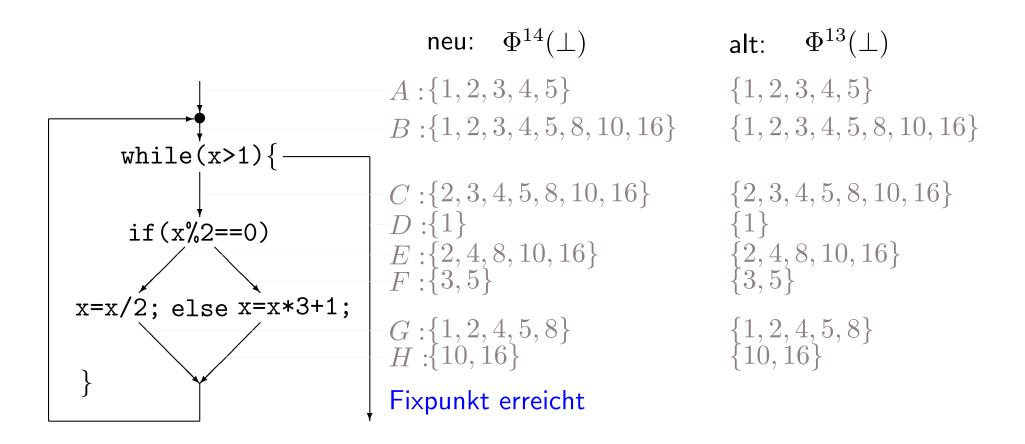


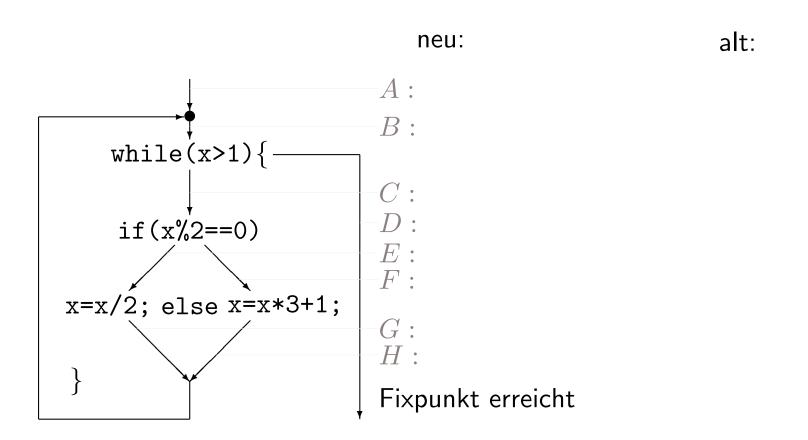


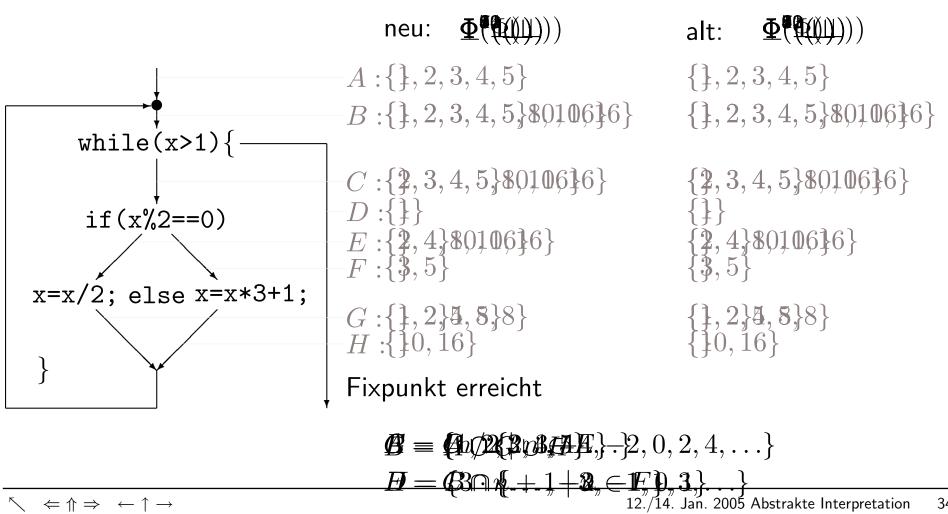










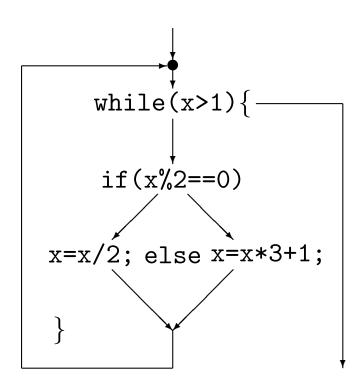


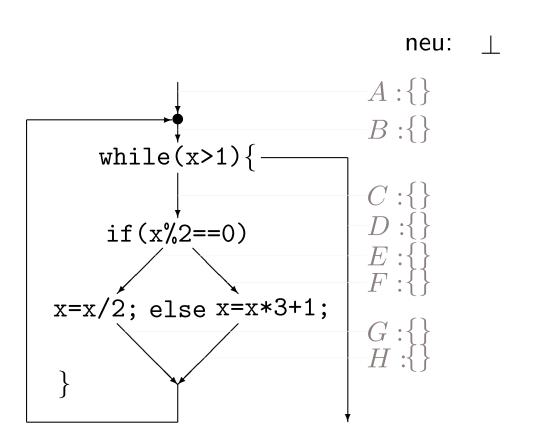
## Fixpunkt erreicht: wieso eigentlich?

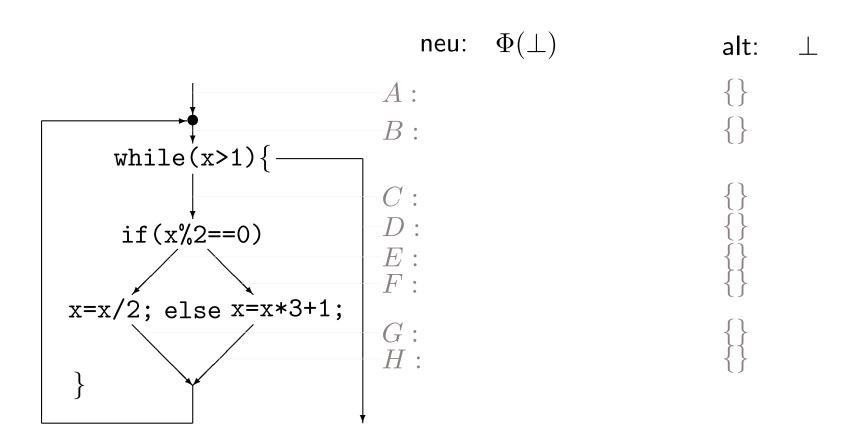
Der kleinste Fixpunkt von  $\Phi$  ist nach dem Fixpunktsatz  $X = \coprod \{\Phi^n(\bot) \mid n \in I\!\!N\}.$ 

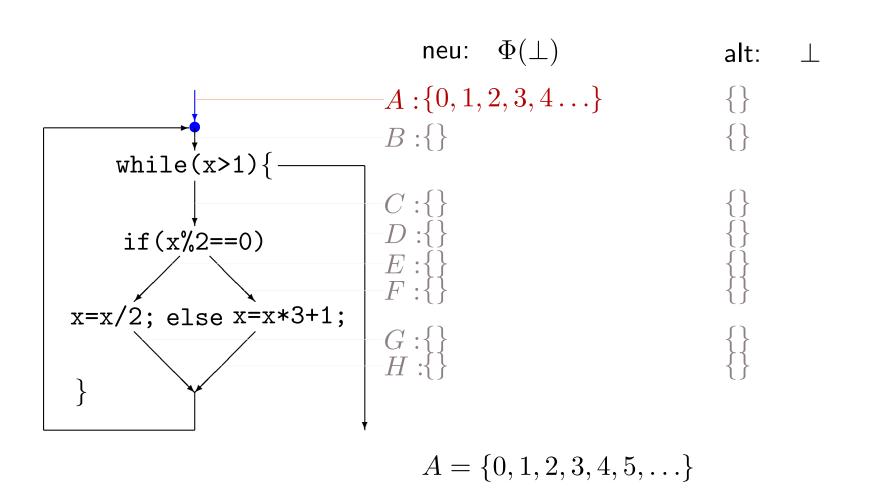
Wir haben gesehen, daß  $\bot \sqsubseteq \Phi(\bot) \sqsubseteq \Phi(\Phi(\bot)) \sqsubseteq \ldots \sqsubseteq \Phi^{13}(\bot) = \Phi^{14}(\bot)$ . Weitere Anwendung von  $\Phi$  bringt keine Änderung mehr, d.h.  $\Phi^{13}(\bot) = \Phi^{14}(\bot) = \ldots = \Phi^n$  für alle n > 14. Daher ist  $\Phi^{13}(\bot)$  die kleinste obere Schranke von  $\{\Phi^n(\bot) \mid n \in I\!\!N\}$ , d.h.  $X = \Phi^{13}(\bot)$ .

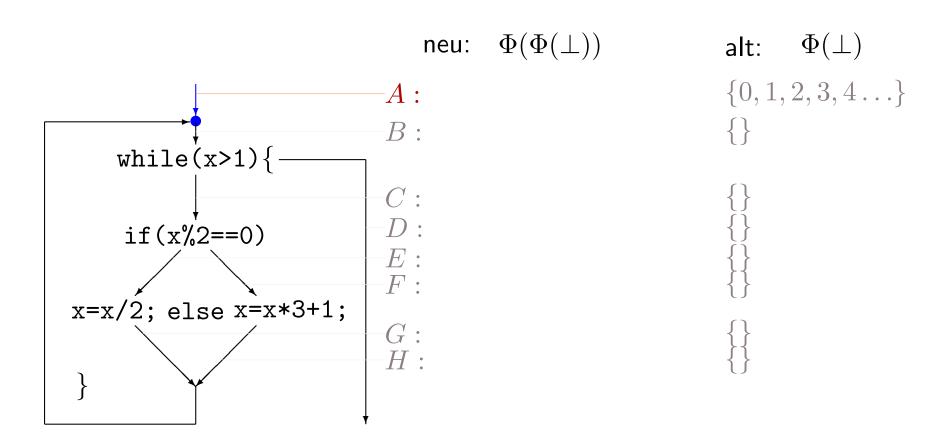
Wenn anstatt  $\{1,2,3,4,5\}$  z.B. ganz  $I\!\!N$  als Menge der Eingabewerte zugelassen ist, können wir die möglichen Variablenwerte durch entsprechende Änderung von  $\Phi$  ebenfalls leicht ausrechnen.

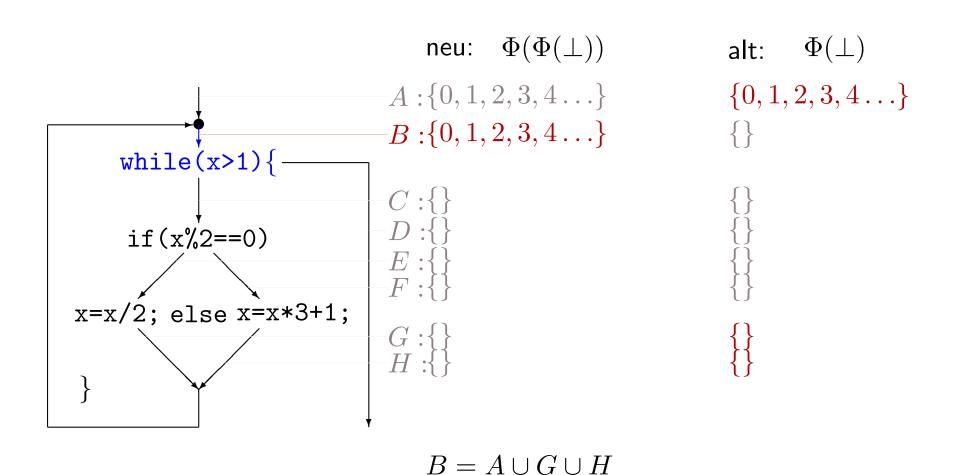


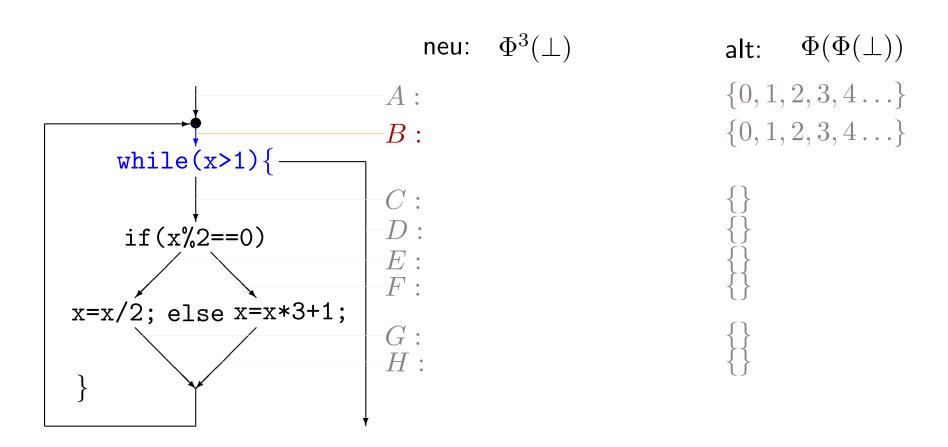


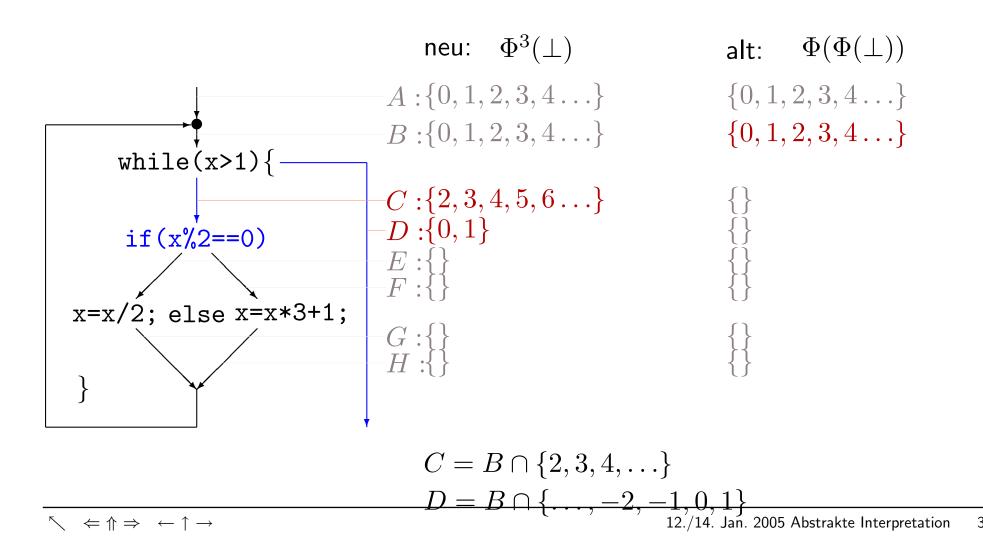


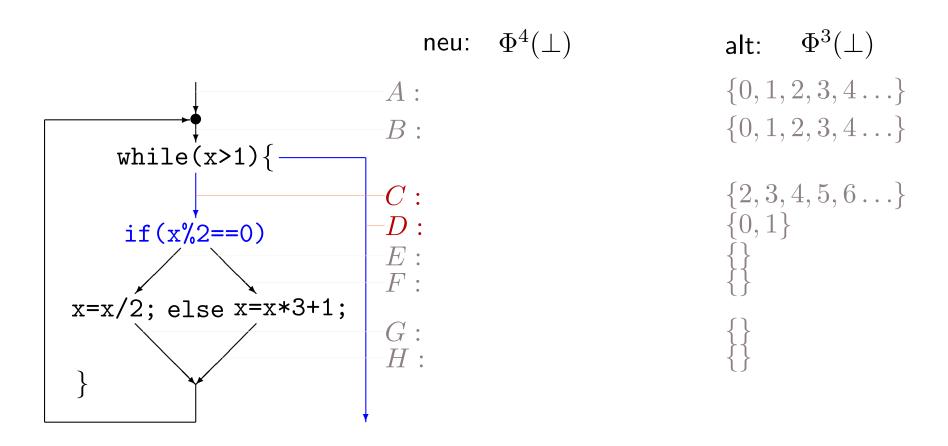


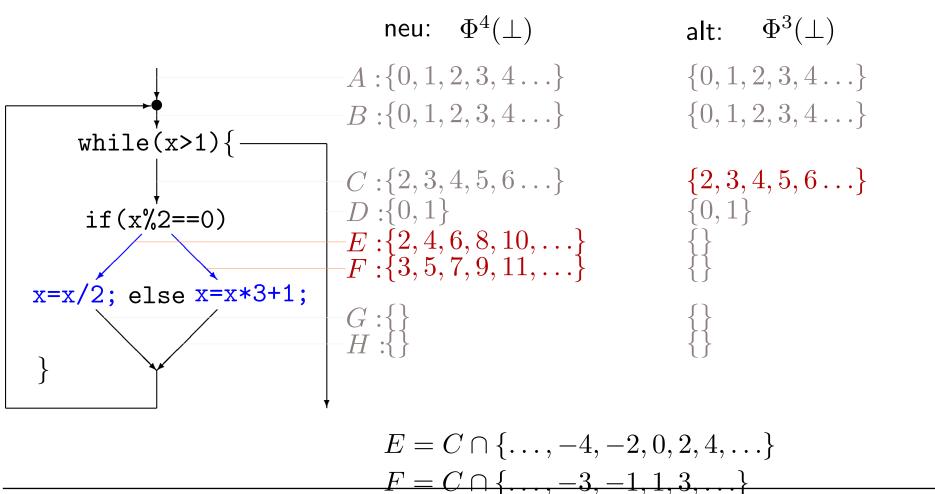


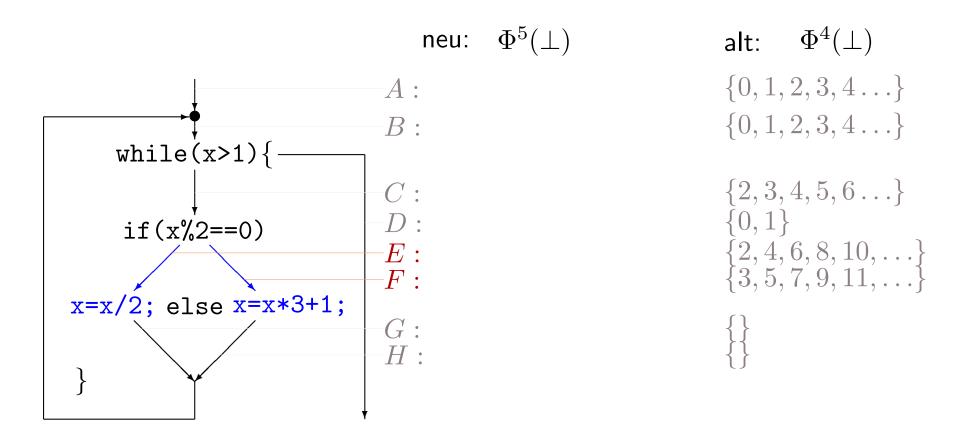


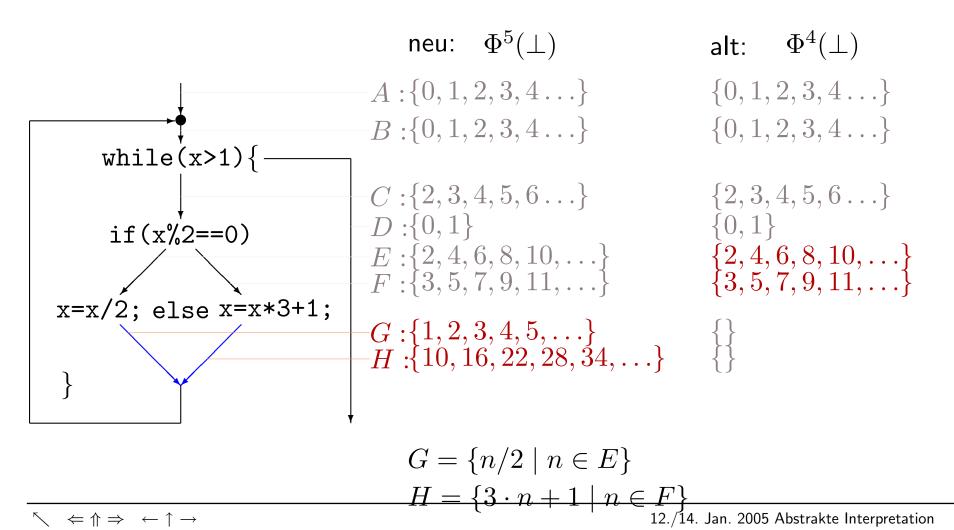


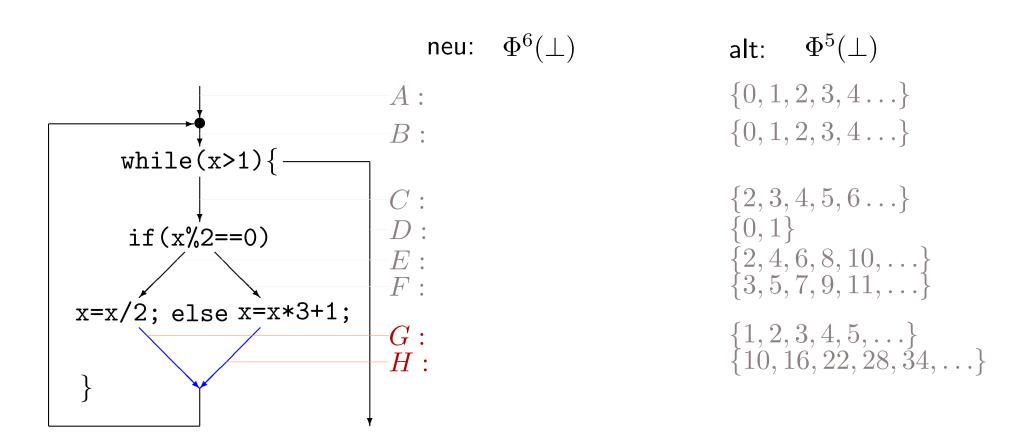


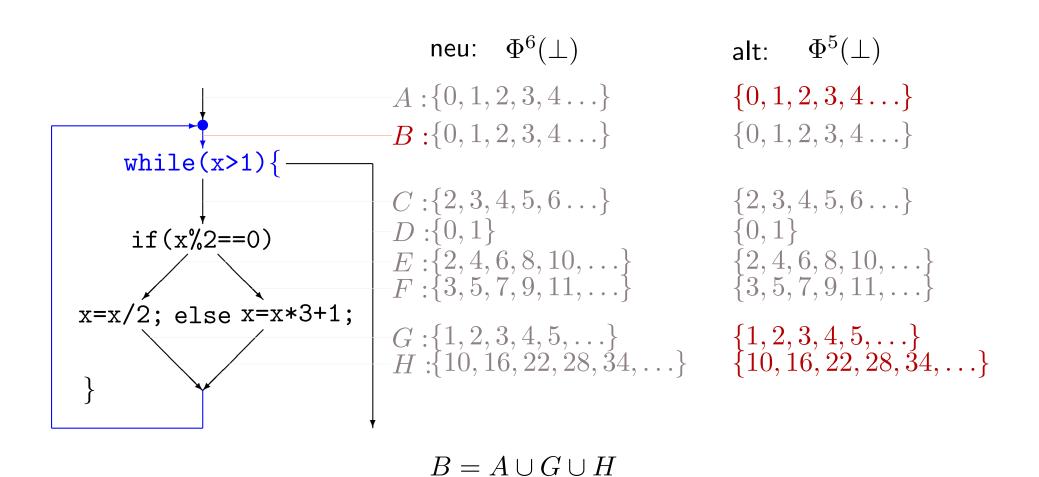


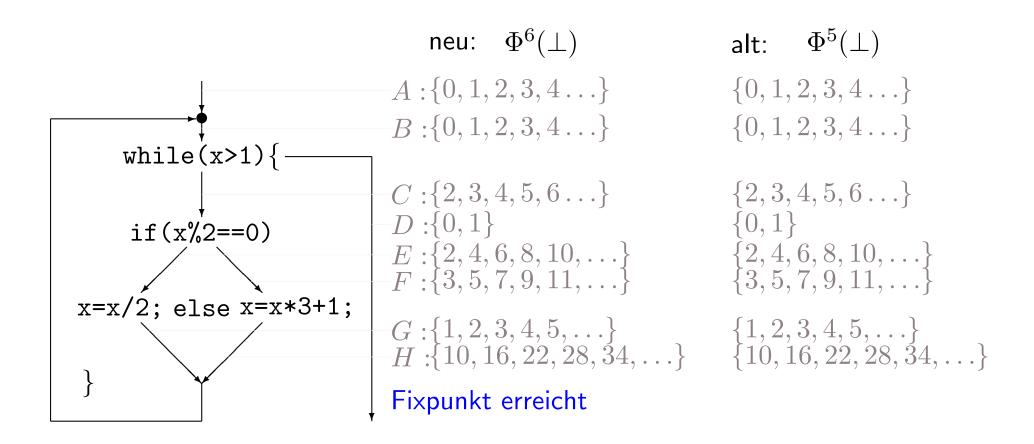


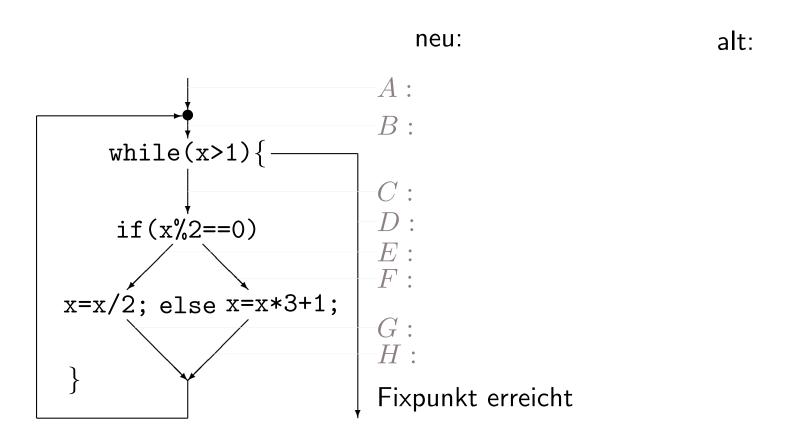


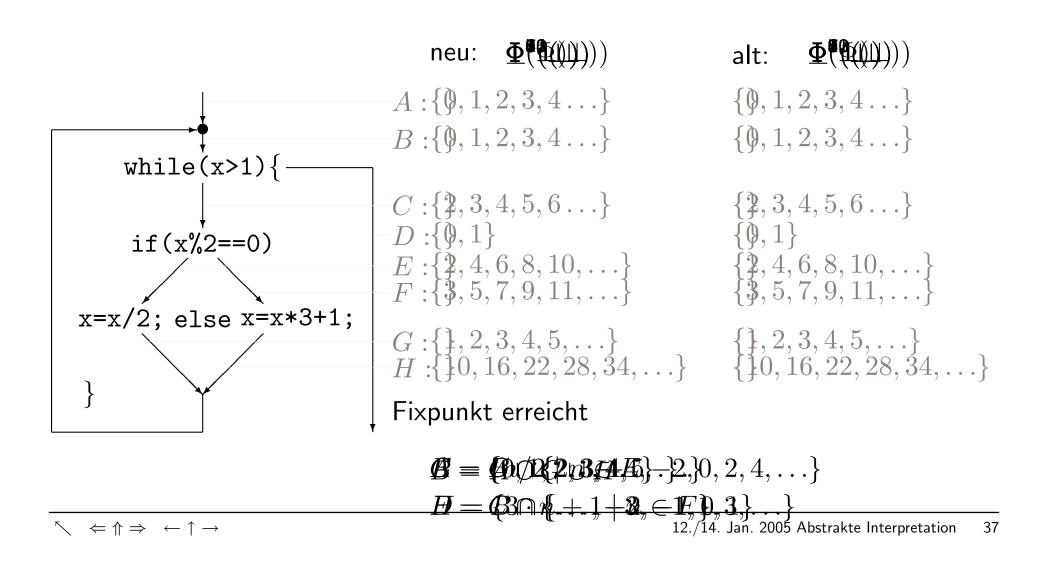






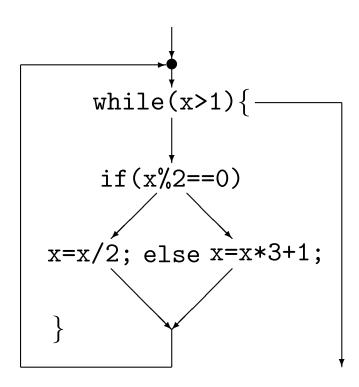


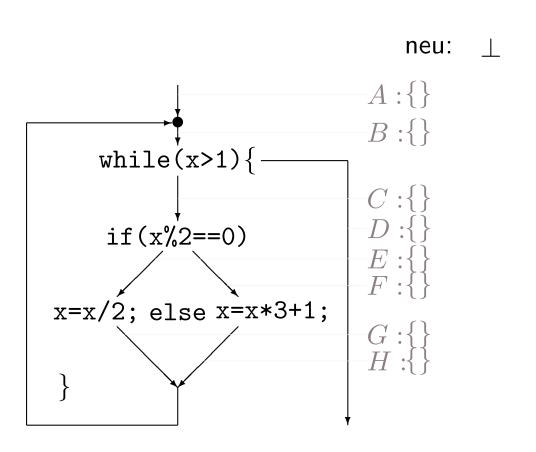


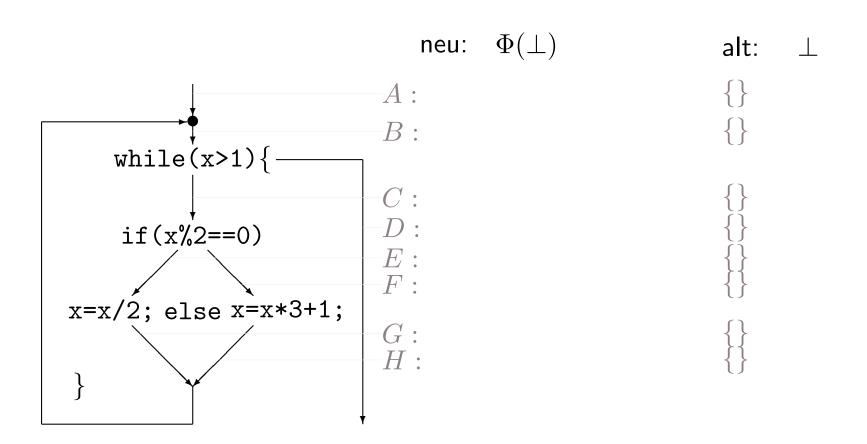


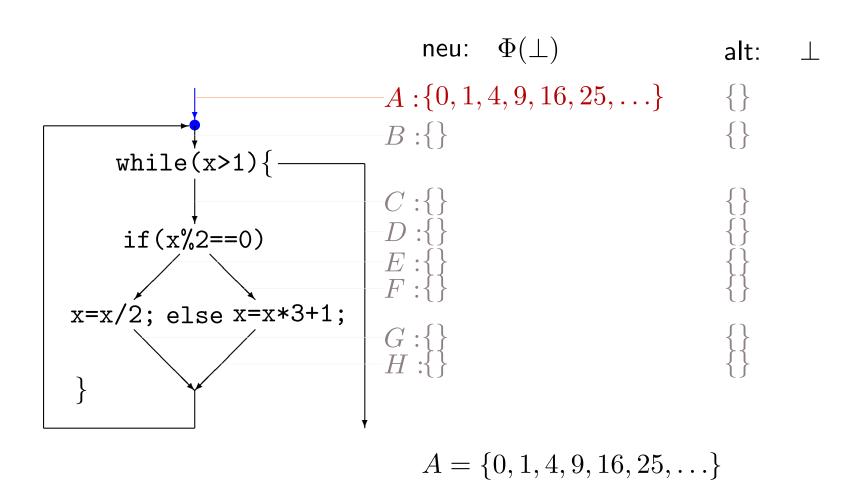
Wenn z.B. die Menge aller Quadratzahlen als Eingabe zugelassen ist (\*), bekommen wir Probleme mit den während der Berechnung auftretenden Mengen.

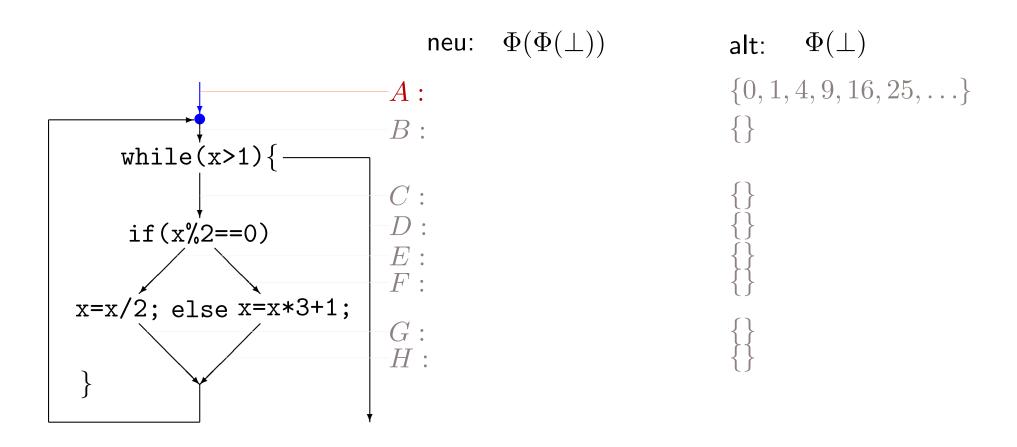
 $^{(*)}$  oder vor das Programm z.B. eine Anweisung x=x\*x; gesetzt wird und davor alle ganzen Zahlen zugelassen werden

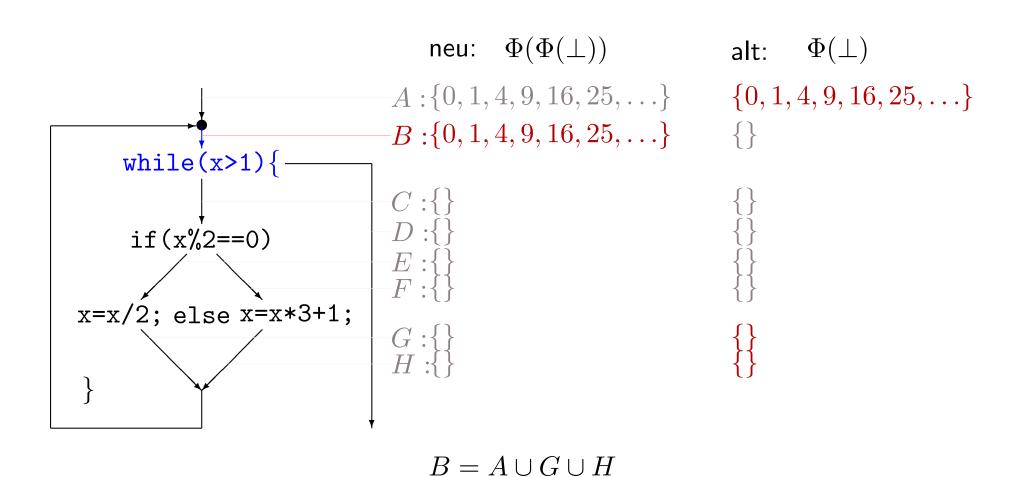


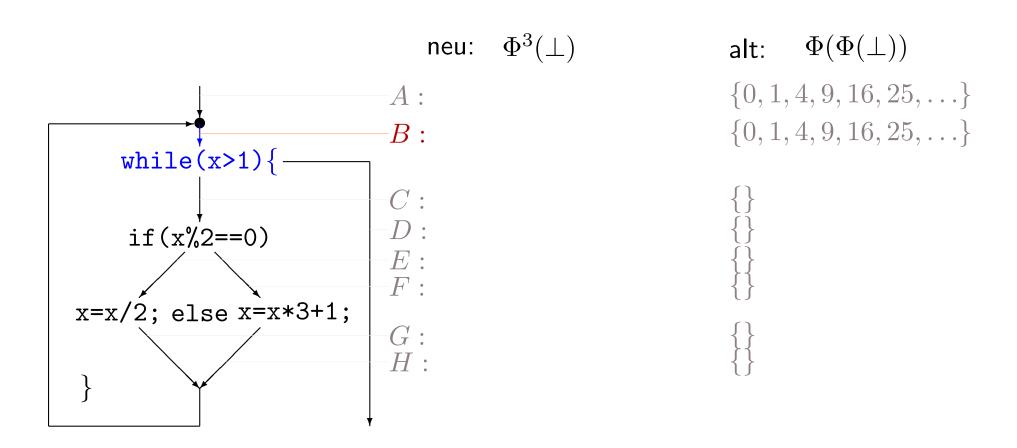


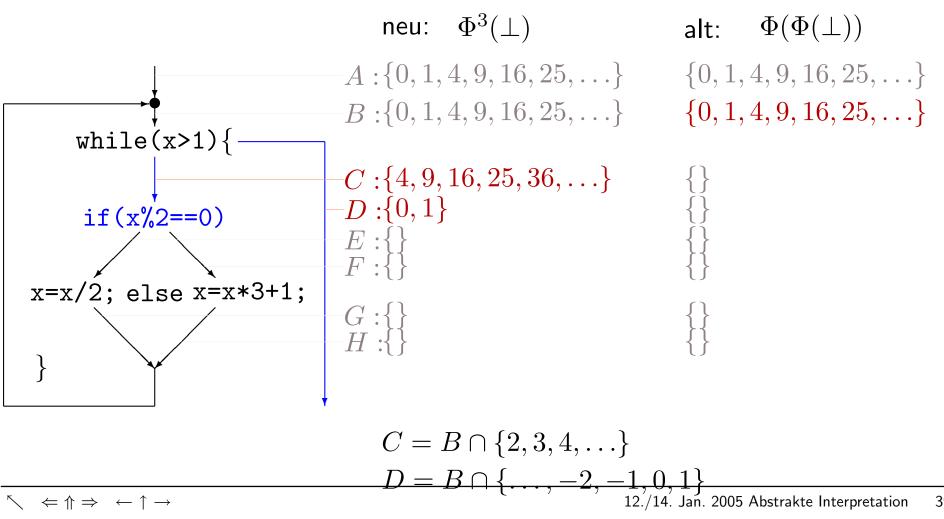


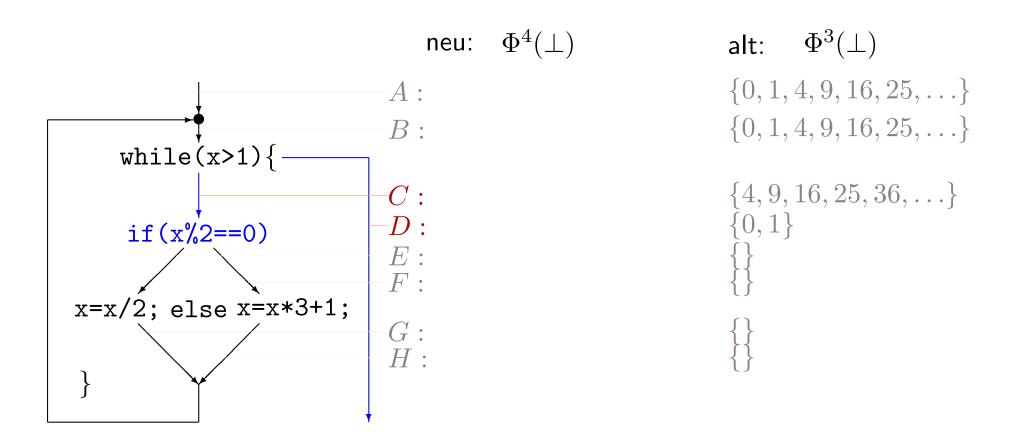


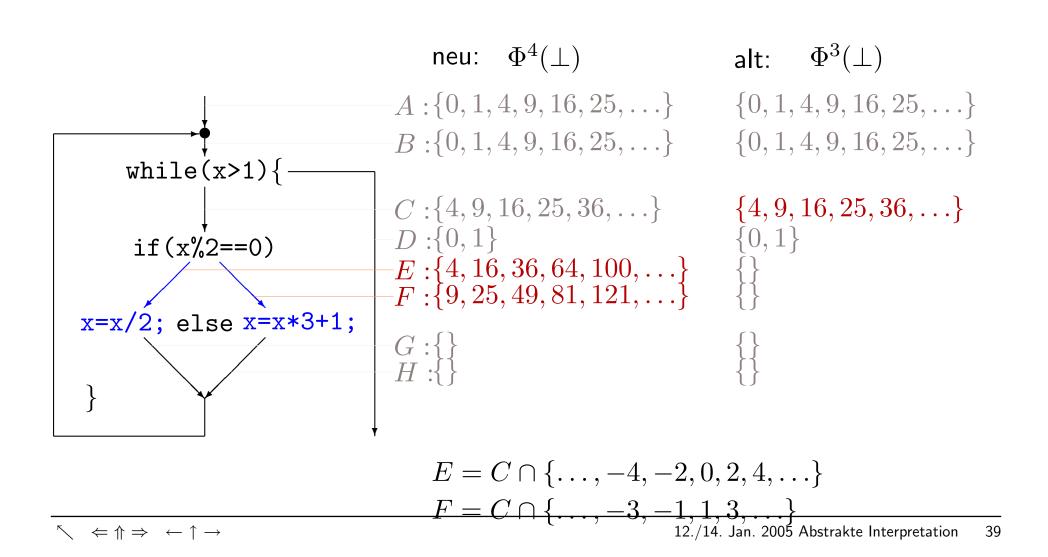


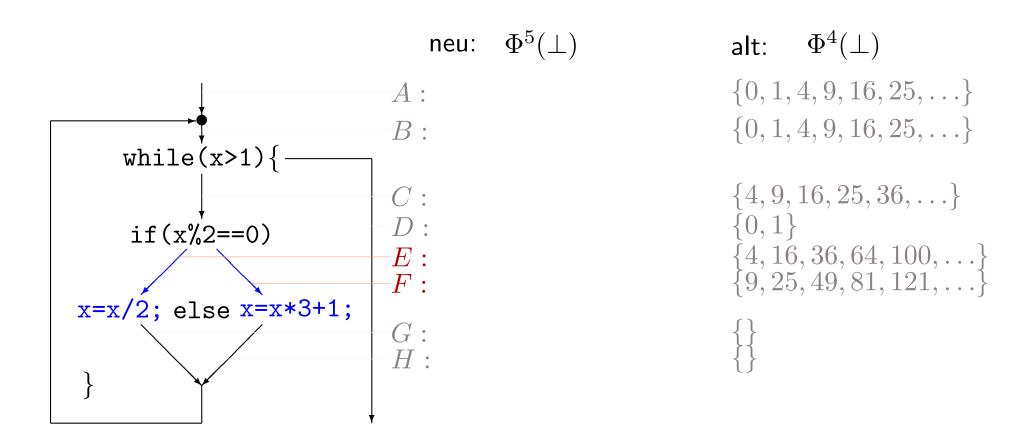


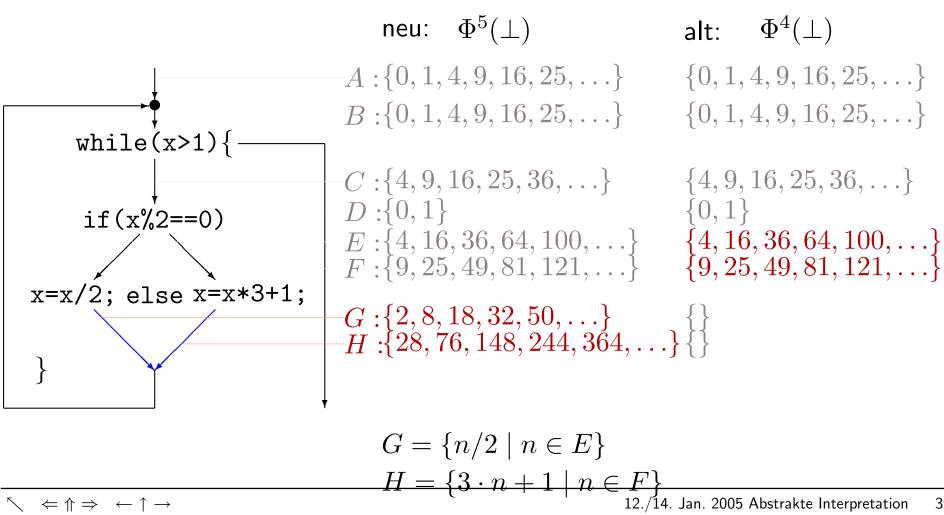


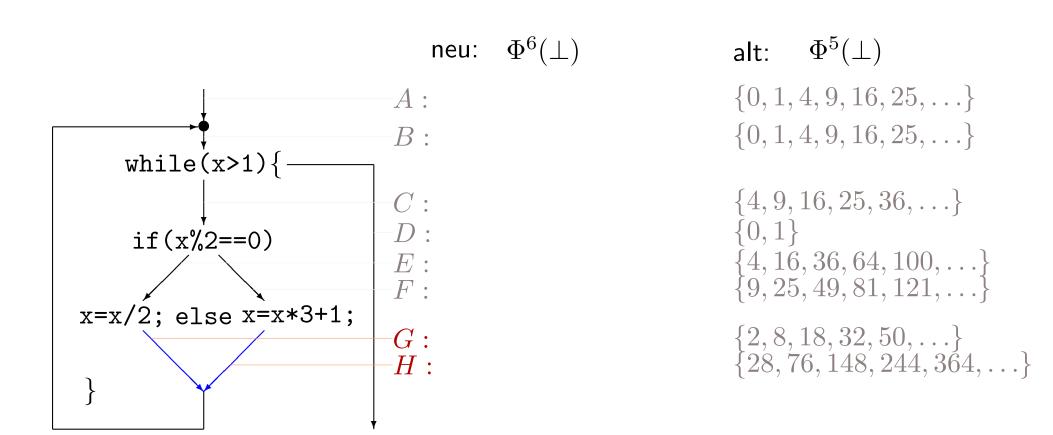


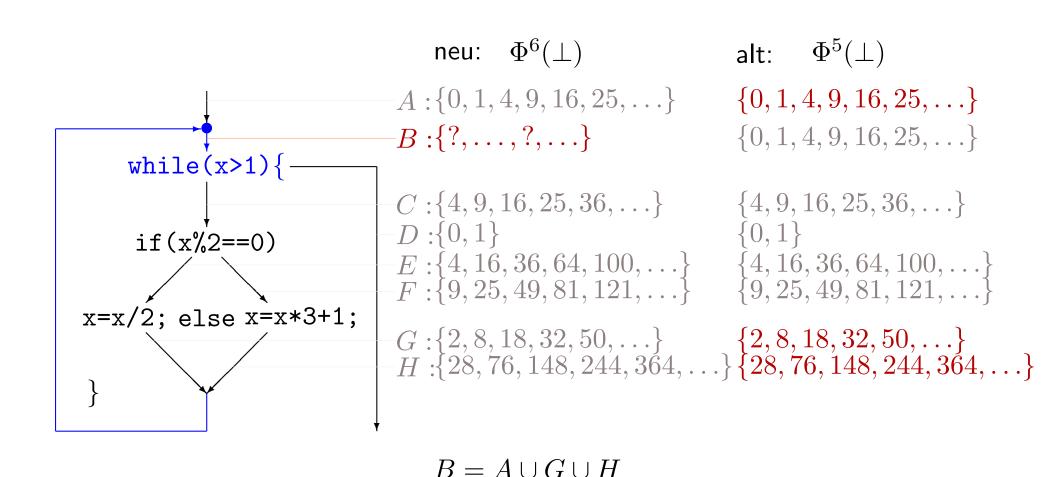


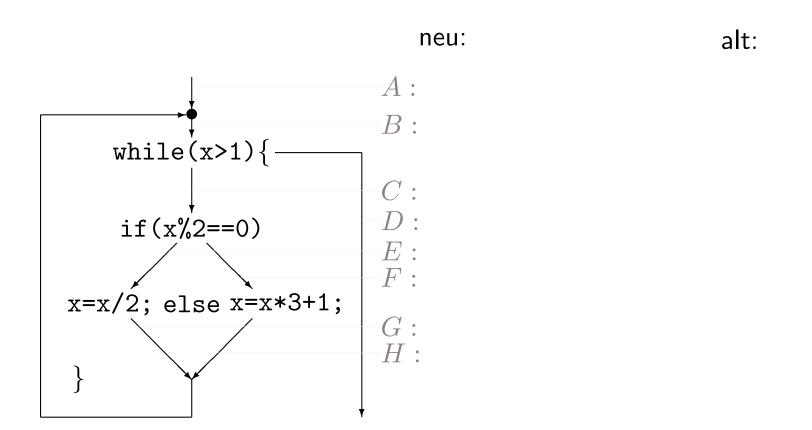


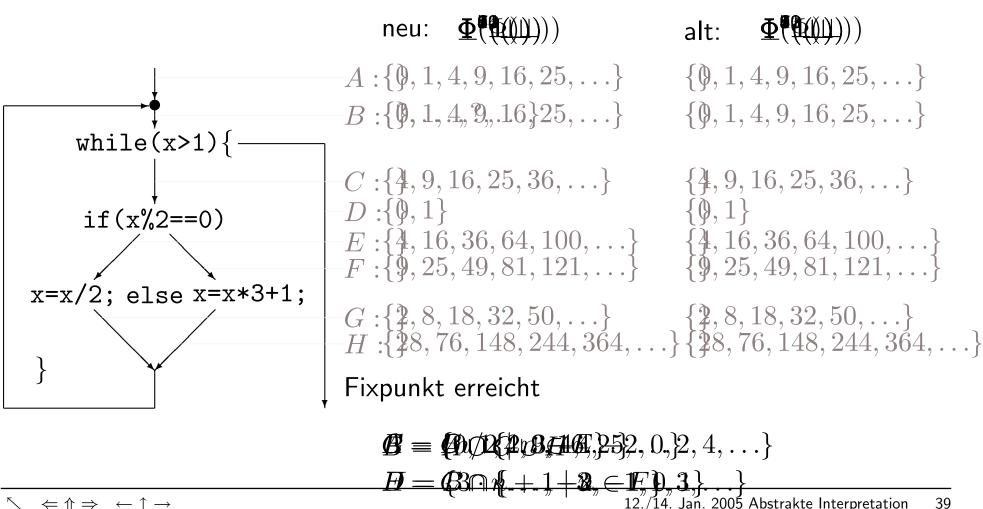












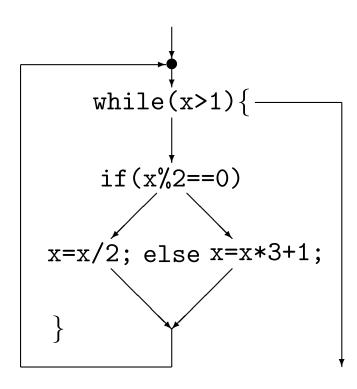
Letztendlich soll die Fixpunktberechnung per Computer durchgeführt werden.

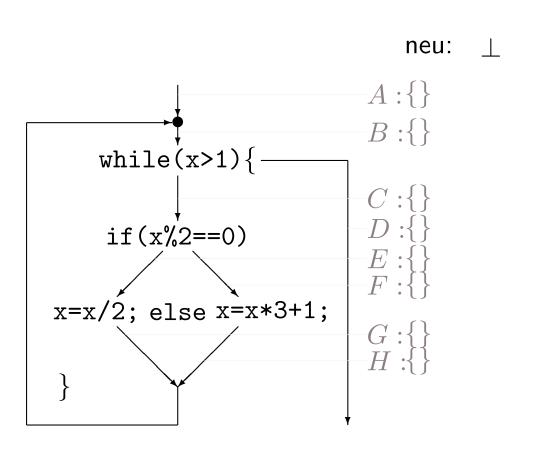
Dafür muß eine geeignete Datenstruktur zur Darstellung der auftretenden Mengen gefunden werden.

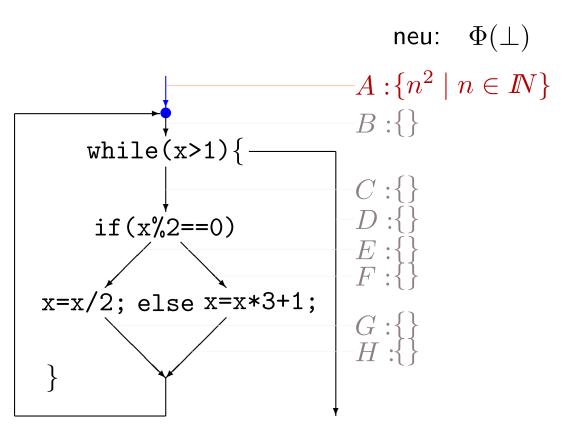
Bitvektoren können nur endliche Mengen (z.B.  $\{0,\ldots,2^{43}\}$  mit 1024 GB Speicher) darstellen.

Treten z.B. 10 Variablen im Programm auf, kann jeweils nur ein Wertebereich  $\{0,\ldots,18\}$  mit 1024 GB dargestellt werden.

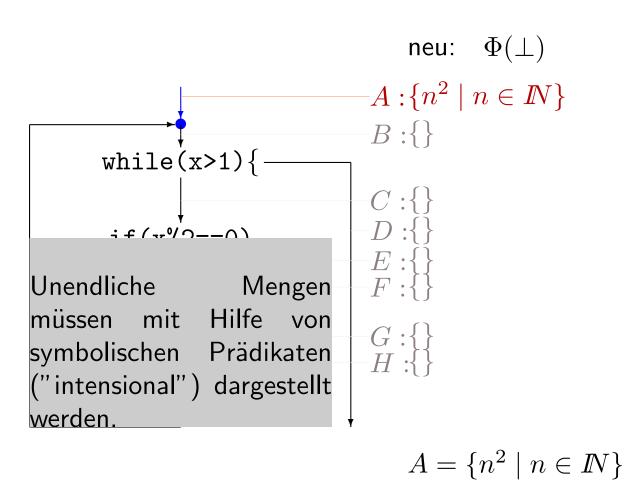
Wir versuchen daher, die Mengen in symbolischer Form darzustellen.

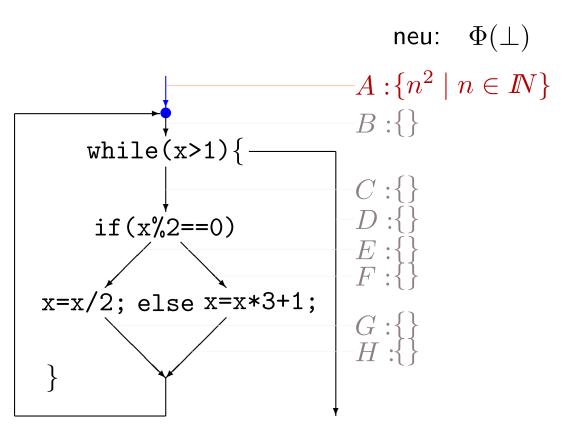




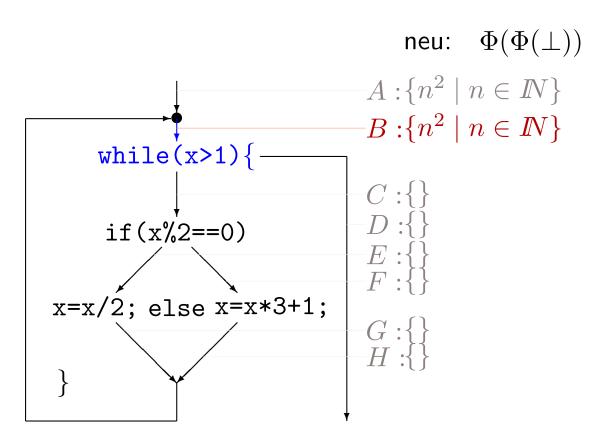


$$A = \{ n^2 \mid n \in \mathbb{N} \}$$

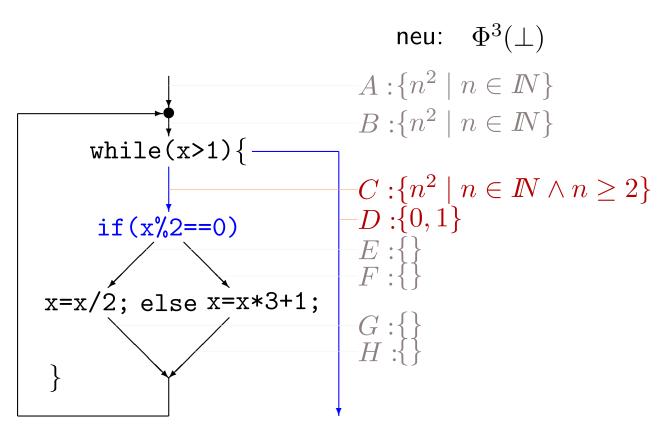




$$A = \{ n^2 \mid n \in \mathbb{N} \}$$

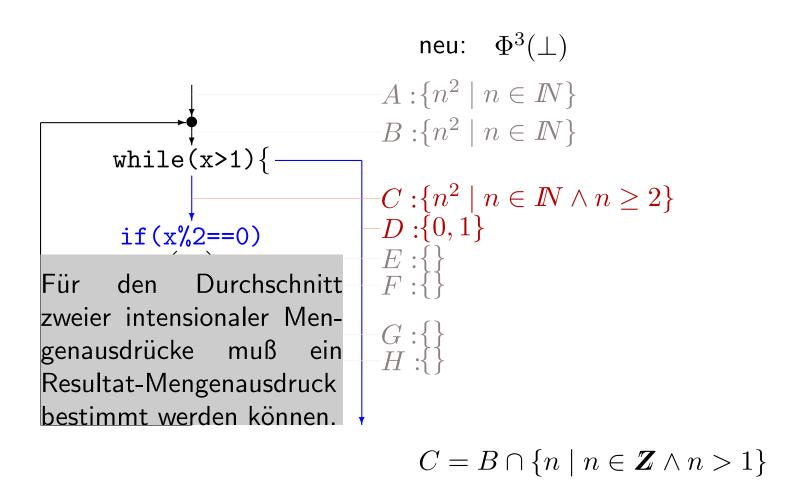


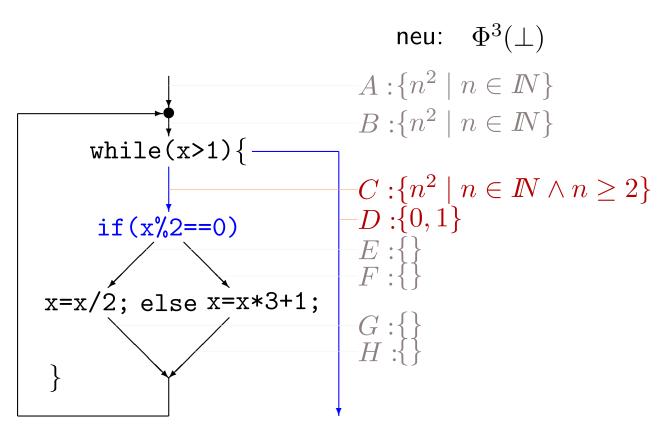
$$B = A \cup G \cup H$$



$$C = B \cap \{n \mid n \in \mathbf{Z} \land n > 1\}$$

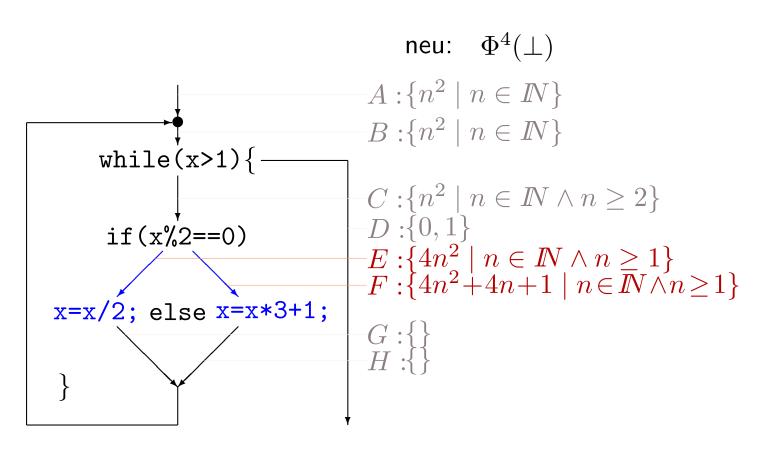
$$D = B \cap \{n \mid n \in \mathbf{Z} \land n \le 1\}$$





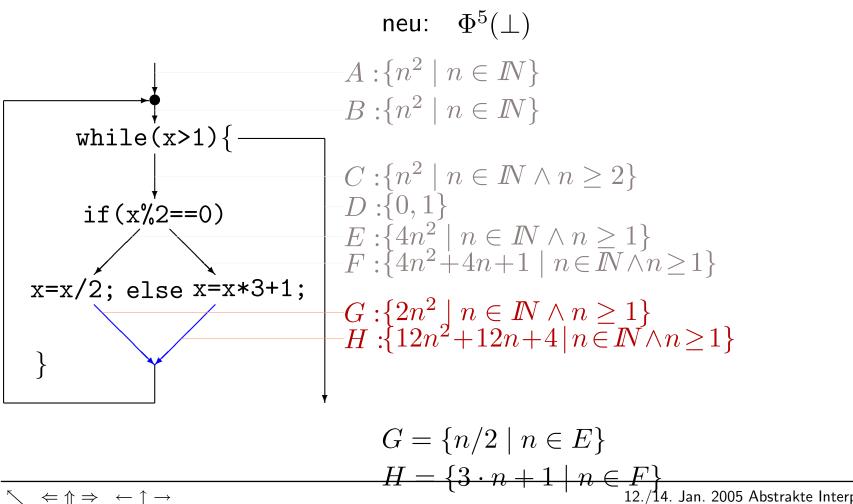
$$C = B \cap \{n \mid n \in \mathbf{Z} \land n > 1\}$$

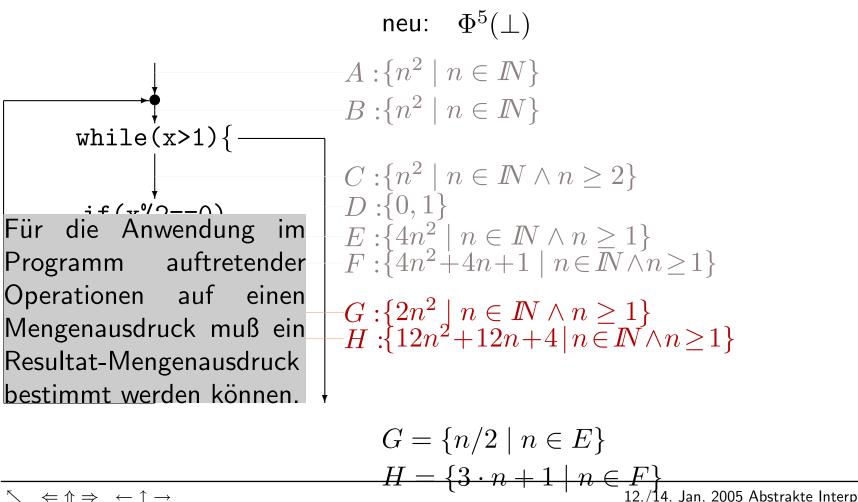
$$D = B \cap \{n \mid n \in \mathbf{Z} \land n \le 1\}$$

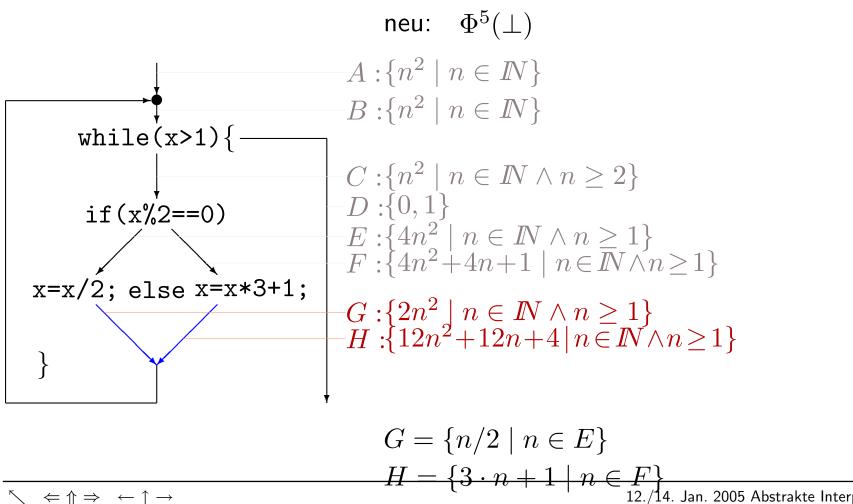


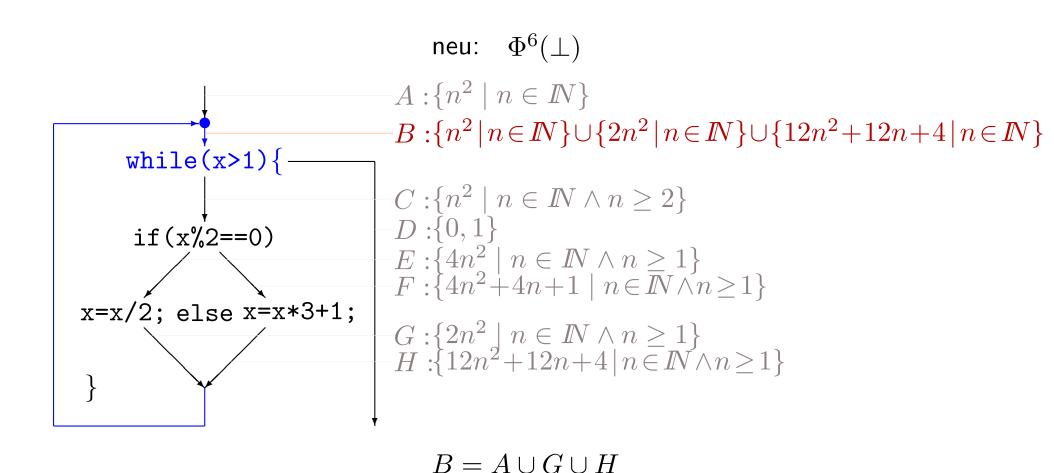
$$E = C \cap \{n \mid n \in \mathbb{Z} \land n\%2 = 0\}$$

$$F = C \cap \{n \mid n \in \mathbb{Z} \land n\%2 \neq 0\}$$
12./14. Jan. 2005 Abstrakte Interpretation









rzhilo(v>1)

Um feststellen zu können, ob ein Fixpunkt erreicht ist, muß entschieden werden können, ob zwei intensionale Mengenausdrücke die gleiche Menge beschreiben.

neu: 
$$\Phi^6(\perp)$$

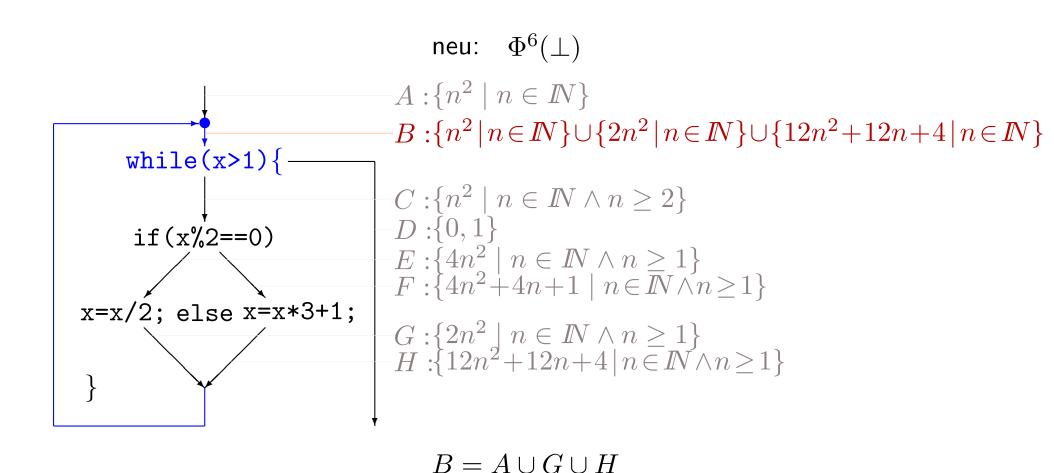
$$A:\{n^2\mid n\in\mathbb{N}\}$$

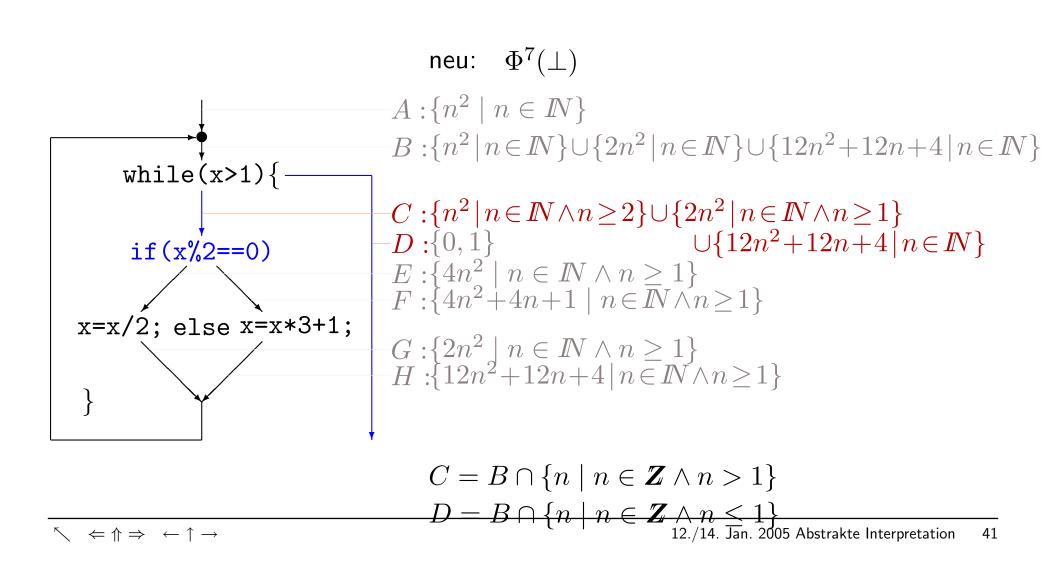
$$B: \{n^2 \mid n \in \mathbb{N}\} \cup \{2n^2 \mid n \in \mathbb{N}\} \cup \{12n^2 + 12n + 4 \mid n \in \mathbb{N}\}$$

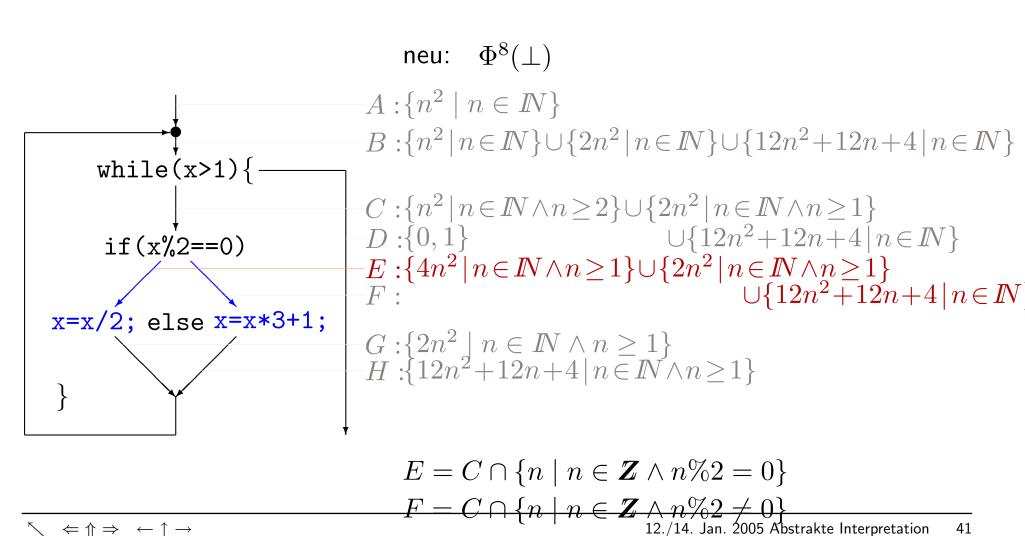
$$\begin{array}{l} C: \{n^2 \mid n \in I\!\!N \land n \geq 2\} \\ D: \{0, 1\} \\ E: \{4n^2 \mid n \in I\!\!N \land n \geq 1\} \\ F: \{4n^2 + 4n + 1 \mid n \in I\!\!N \land n \geq 1\} \end{array}$$

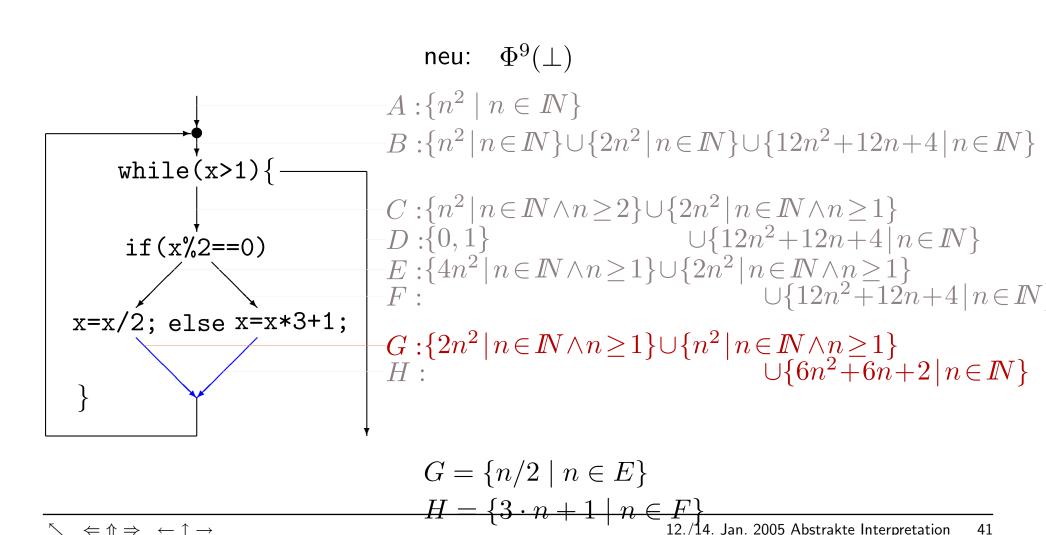
$$G: \{2n^2 \mid n \in \mathbb{N} \land n \ge 1\}$$
  
 $H: \{12n^2 + 12n + 4 \mid n \in \mathbb{N} \land n \ge 1\}$ 

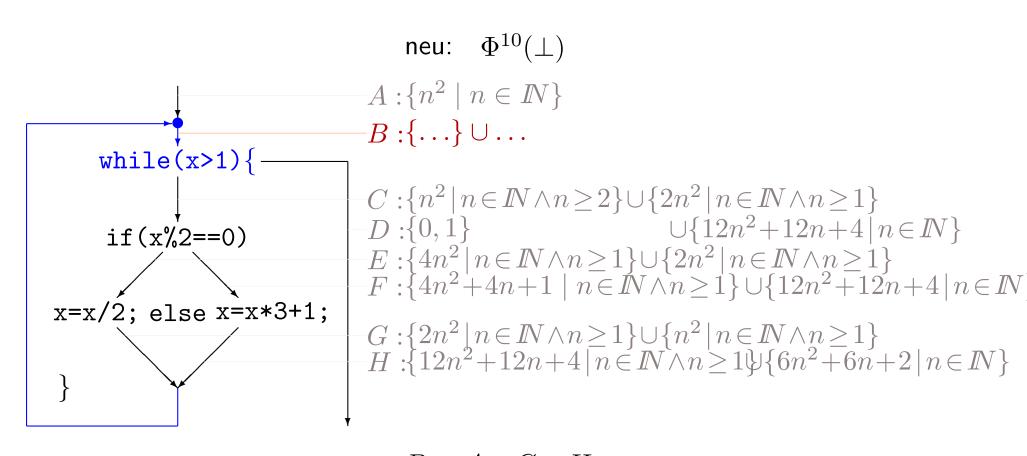
$$B = A \cup G \cup H$$

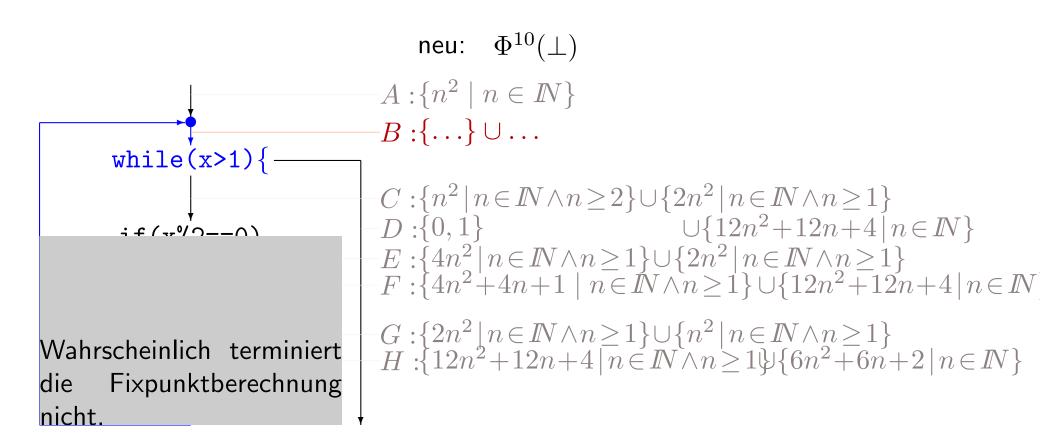




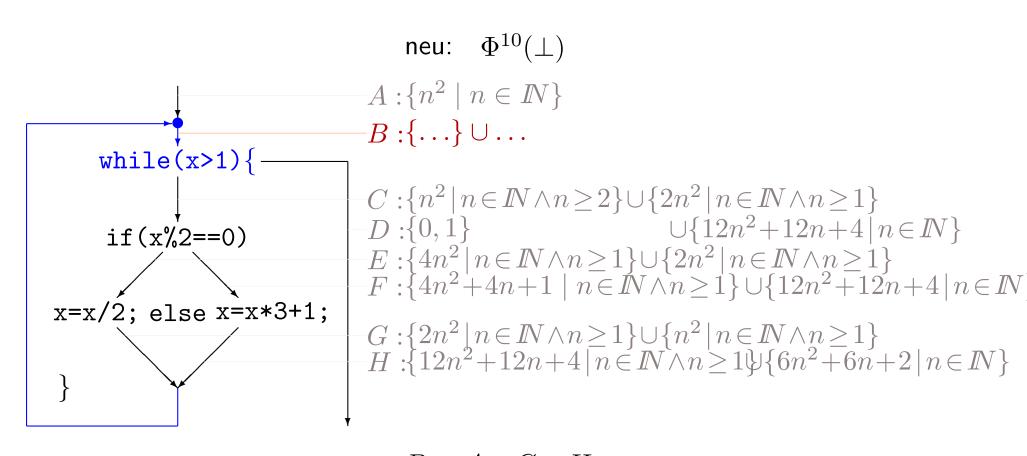


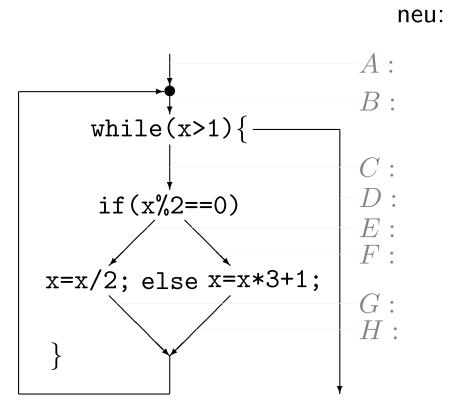






$$B = A \cup G \cup H$$





neu:  $\Phi((\mathbb{M}))$ 

 $A:\{n^2\mid n\in\mathbb{N}\}$  $B: \{n^2 \mid n \in \mathbb{N}\} \cup \{12n^2 \mid n \in \mathbb{N}\} \cup \{12n^2 + 12n + 4 \mid n \in \mathbb{N}\}$ 

Um feststellen zu können,

Wahrscheinlich terminiert Fixpunktberechnung die nicht.

```
C: \{n^2 \mid n \in IN \land n \geq 1\}
       \cup \{12n^2 + 12n + 4 \mid n \in \mathbb{N}\}\
D:\{\emptyset,1\}
```

Fixpunkt erreicht

#### Rechnen mit symbolischen intensionalen Mengenausdrücken

#### Zusammenfassung der Anforderungen:

- 1. Mindestens Durchschnitt, Vereinigung und Programmoperations-Anwendung muß für Mengenausdrücke berechenbar sein, das Resultat muß insbesondere ausdrückbar sein.
- 2. Die (extensionale) Gleichheit von Mengenausdrücken muß entscheidbar sein.

#### Dafür gibt es keine geeigneten Kalküle.

Wegen 2. muß die Form der in Mengenausdrücken erlaubten Prädikate stark eingeschränkt werden.

Dann können aber Mengen, die sich aus der Anwendung von Programmoperationen ergeben, nur in den seltensten Fällen noch ausgedrückt werden.

Außerdem können wir nicht garantieren, daß die Fixpunktberechnung terminiert.

#### ABSTRAKTE INTERPRETATION

Die genaue Menge der möglichen Variablenwerte an jedem Programmpunkt läßt sich also i.allg. nicht automatisch berechnen.

Wir versuchen daher im Folgenden, wenigstens grobe Informationen über die Programmpunkte zu bestimmen.

Dazu fassen wir jeweils mehrere Einzelwerte zu einem "abstrakten Wert" zusammen.

### **Abstrakte Interpretation**

Je nach Anwendung kommen z.B. infrage:

| konkrete Werte  | abstrakter Wert |
|-----------------|-----------------|
| , -3, -2, -1    | neg             |
| 0               | null            |
| 1, 2, 3,        | pos             |
| 4, -2, 0, 2, 4, | even            |
| 3, -1, 1, 3,    | odd             |

#### **Abstrakte Interpretation**

Da wir ohnehin mit Wertemengen statt Einzelwerten rechnen, können wir auch Mengen abstrahieren:

| konkrete Mengen   | abstrakte Menge      |
|---|----------------------|
| $\{4,8\}, \{4,7,8\}, \{4,5,6,7,8\}, \dots$  | $\boxed{[4\dots 8]}$ |
| $\{1, 2, 3, 4, \ldots\}, \{1, 2, 4, 8, \ldots\}, \{1, 4, 9, 16, \ldots\}, \ldots$ | $[1\dots\infty]$     |

Wenn wir zu jedem Programmpunkt die abstrakte Menge bestimmen können, können wir wenigstens Aussagen über bestimmte Eigenschaften machen, je nach Abstraktionsverfahren z.B. über Vorzeichen oder Wertebereich einer Variablen.

#### Anforderungen an eine Abstraktion

#### Korrektheit:

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

#### Uberführbarkeit:

Jedes Programmkonstrukt (Verzweigung, Zusammenführung, Zuweisung, . . . ) muß sich in eine abstrakte Operation / Gleichung überführen lassen. Insbesondere die im Programm auftretenden Operationen müssen eine abstrakte Entsprechung haben.

#### Terminierung:

Wir suchen ein Verfahren zur Berechnung der abstrakten Mengen, von dem wir garantieren können, daß es immer terminiert.

#### Aussagekraft:

Die abstrakten Mengen sollen genügend präzise sein, um die (benutzergegebenen) Eigenschaften an den Programmpunkten entscheiden zu können.

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

```
Satz (Cousot, Cousot 1976):

Sei M mit (\sqsubseteq) und M' mit (\sqsubseteq') vollständiger Verband,

M \stackrel{\gamma}{\rightleftharpoons} M' eine Galois-Verbindung,

\Phi: M \longrightarrow M monoton und stetig,

X der kleinste Fixpunkt von \Phi und

X' = \bigsqcup' \{ \bot', \ \Phi'(\bot'), \ \Phi'(\Phi'(\bot')), \ \Phi'(\Phi'(\bot')), \ \ldots \}.

Dann ist X \sqsubseteq \gamma(X').
```

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

```
Satz (Cousot, Cousot 1976):

Sei M mit (\sqsubseteq) und M' mit (\sqsubseteq') vollständiger Verband,

M \stackrel{?}{\rightleftharpoons} M' eine Galois-Verbindung,

\Phi : M \longrightarrow M monoton und stetig,

X der kleinste Fixpunkt von \Phi und

X' = \bigsqcup' \{ \bot', \ \Phi'(\bot'), \ \Phi'(\Phi'(\bot')), \ \Phi'(\Phi'(\bot')), \ \ldots \}.

Dann ist X \sqsubseteq \gamma(X').
```

Wir benötigen noch eine weitere formale Definition.

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

```
Satz (Cousot, Cousot 1976): 
Sei M mit (\sqsubseteq) und M' mit (\sqsubseteq') vollständiger Verband, M \stackrel{\gamma}{=} M' eine Galois-Verbindung, \Phi: M \longrightarrow M monoton und stetig, X der kleinste Fixpunkt von \Phi und X' = \bigsqcup' \{ \bot', \; \Phi'(\bot'), \; \Phi'(\Phi'(\bot')), \; \Phi'(\Phi'(\bot')), \; \ldots \}. Dann ist X \sqsubseteq \gamma(X').
```

Die für einen Programmpunkt berechnete abstrakte Wertemenge soll die tatsächlich auftretende konkrete Wertemenge nach oben abschätzen.

```
Satz (Cousot, Cousot 1976):

Sei M mit (\sqsubseteq) und M' mit (\sqsubseteq') vollständiger Verband, M \overset{\gamma}{\underset{\alpha}{\longrightarrow}} M' eine Galois-Verbindung, \Phi: M \longrightarrow M monoton und stetig, X der kleinste Fixpunkt von \Phi und X' = \bigsqcup' \{ \bot', \ \Phi'(\bot'), \ \Phi'(\Phi'(\bot')), \ \Phi'(\Phi'(\bot')), \ \ldots \}. Dann ist X \sqsubseteq \gamma(X').
```

Wir benötigen noch eine weitere formale Definition.

### **Galois-Verbindung**

Seien M mit  $(\sqsubseteq)$  und M' mit  $(\sqsubseteq')$  vollständige Verbände.

Sei  $\alpha: M \longrightarrow M'$  und  $\gamma: M' \longrightarrow M$  monotone Abbildungen, so daß  $x \sqsubseteq \gamma(\alpha(x))$  für alle  $x \in M$  und  $\alpha(\gamma(x')) \sqsubseteq' x'$  für alle  $x' \in M'$ .

Wir nennen das Paar  $\alpha, \gamma$  eine Galois-Verbindung zwischen M und M' und schreiben  $M \stackrel{\gamma}{\underset{\alpha}{\rightleftharpoons}} M'$ .

 $\alpha(x)$  ist die Abstraktion von x, d.h. die präzisest-mögliche Darstellung von x in M'.

 $\gamma(x')$  ist die Konkretisierung von x', d.h. das unpräziseste Element in M, das noch durch x' dargestellt werden kann.

 $M=\wp(Z)$  mit  $(\subseteq)$ ,  $M'=\wp(\{-,0,+\})$  mit  $(\subseteq)$ . Wir definieren die Abstraktion  $\alpha$  Für beliebige Mengen zunächst auf einelementigen Mengen:  $S\subseteq Z$  definieren wir:

$$\alpha(\{n\}) = \begin{cases} \{-\} & \text{falls} \quad n < 0 \\ \{0\} & \text{falls} \quad n = 0 \\ \{+\} & \text{falls} \quad n > 0 \end{cases} \qquad \alpha(S) = \bigcup_{n \in S} \alpha(\{n\}).$$

 $M = \wp(\mathbf{Z}) \text{ mit } (\subseteq), \quad M' = \wp(\{-,0,+\}) \text{ mit } (\subseteq).$  Wir definieren die Abstraktion  $\alpha$  Für beliebige Mengen zunächst auf einelementigen Mengen:  $S \subseteq \mathbf{Z}$  definieren wir:

$$\alpha(\{\mathbf{n}\}) = \begin{cases} \{-\} & \text{falls} \quad \mathbf{n} < 0 \\ \{0\} & \text{falls} \quad \mathbf{n} = 0 \\ \{+\} & \text{falls} \quad \mathbf{n} > 0 \end{cases} \qquad \alpha(S) = \bigcup_{\mathbf{n} \in S} \alpha(\{\mathbf{n}\}).$$

Z.B.

$$\alpha(\{-2,-1\}) = \alpha(\{-2\}) \cup \alpha(\{-1\}) = \{-\} \cup \{-\} = \{-\},\$$

$$\alpha(\{0,2,4,6,\ldots\}) = \alpha(\{0\}) \cup \alpha(\{2\}) \cup \alpha(\{4\}) \cup \alpha(\{6\}) \cup \ldots = \{0,+\}.$$

 $M = \wp(\mathbf{Z}) \text{ mit } (\subseteq), \quad M' = \wp(\{-,0,+\}) \text{ mit } (\subseteq).$ 

Wir definieren die Abstraktion  $\alpha$  Für beliebige Mengen

zunächst auf einelementigen Mengen:  $S \subseteq \mathbb{Z}$  definieren wir:

$$\alpha(\{n\}) = \begin{cases} \{-\} & \text{falls} \quad n < 0 \\ \{0\} & \text{falls} \quad n = 0 \\ \{+\} & \text{falls} \quad n > 0 \end{cases} \qquad \alpha(S) = \bigcup_{n \in S} \alpha(\{n\}).$$

Entsprechend definieren wir die Konkretisierung  $\gamma$  durch

$$\gamma(\{-\}) = \{\dots, -3, -2, -1\} 
\gamma(\{0\}) = \{0\} 
\gamma(\{+\}) = \{1, 2, 3, \dots\} 
\gamma(S') = \bigcup_{s \in S'} \gamma(\{s\}).$$

 $M = \wp(\mathbf{Z}) \text{ mit } (\subseteq), \quad M' = \wp(\{-,0,+\}) \text{ mit } (\subseteq).$ 

Wir definieren die Abstraktion lpha Für beliebige Mengen

zunächst auf einelementigen Mengen:  $S \subseteq \mathbb{Z}$  definieren wir:

$$\alpha(\{\mathbf{n}\}) = \begin{cases} \{-\} & \text{falls} \quad \mathbf{n} < 0 \\ \{0\} & \text{falls} \quad \mathbf{n} = 0 \\ \{+\} & \text{falls} \quad \mathbf{n} > 0 \end{cases} \qquad \alpha(\mathbf{S}) = \bigcup_{\mathbf{n} \in \mathbf{S}} \alpha(\{\mathbf{n}\}).$$

Entsprechend definieren wir die Konkretisierung  $\gamma$  durch

$$\gamma(\{-\}) = \{\dots, -3, -2, -1\} 
\gamma(\{0\}) = \{0\} 
\gamma(\{+\}) = \{1, 2, 3, \dots\} 
\gamma(S') = \bigcup_{s \in S'} \gamma(\{s\}).$$

Z.B.

$$\gamma(\{-,0,+\}) = \gamma(\{-\}) \cup \gamma(\{0\}) \cup \gamma(\{+\})$$
  
=  $\{\ldots,-3,-2,-1\} \cup \{0\} \cup \{1,2,3,\ldots\} = \mathbf{Z}.$ 

 $M = \wp(\mathbf{Z}) \text{ mit } (\subseteq), \quad M' = \wp(\{-,0,+\}) \text{ mit } (\subseteq).$ 

Wir definieren die Abstraktion  $\alpha$  Für beliebige Mengen

zunächst auf einelementigen Mengen:  $S \subseteq \mathbb{Z}$  definieren wir:

$$\alpha(\{n\}) = \begin{cases} \{-\} & \text{falls} \quad n < 0 \\ \{0\} & \text{falls} \quad n = 0 \\ \{+\} & \text{falls} \quad n > 0 \end{cases} \qquad \alpha(S) = \bigcup_{n \in S} \alpha(\{n\}).$$

Entsprechend definieren wir die Konkretisierung  $\gamma$  durch

$$\gamma(\{-\}) = \{\dots, -3, -2, -1\} 
\gamma(\{0\}) = \{0\} 
\gamma(\{+\}) = \{1, 2, 3, \dots\} 
\gamma(S') = \bigcup_{s \in S'} \gamma(\{s\}).$$

 $M = \wp(\mathbf{Z}) \text{ mit } (\subseteq), \quad M' = \wp(\{-,0,+\}) \text{ mit } (\subseteq).$ 

Wir definieren die Abstraktion lpha Für beliebige Mengen

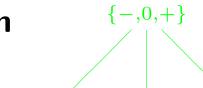
zunächst auf einelementigen Mengen:  $S \subseteq \mathbb{Z}$  definieren wir:

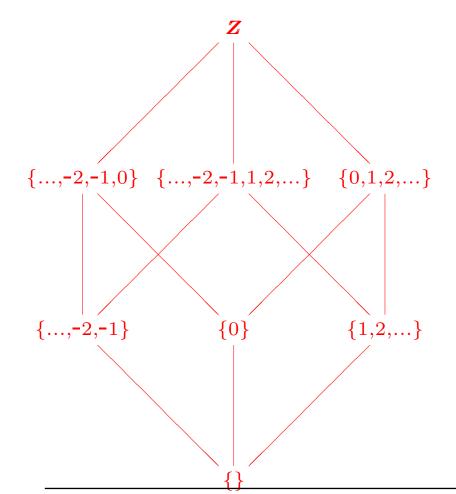
$$\alpha(\{n\}) = \begin{cases} \{-\} & \text{falls} \quad n < 0 \\ \{0\} & \text{falls} \quad n = 0 \\ \{+\} & \text{falls} \quad n > 0 \end{cases} \qquad \alpha(S) = \bigcup_{n \in S} \alpha(\{n\}).$$

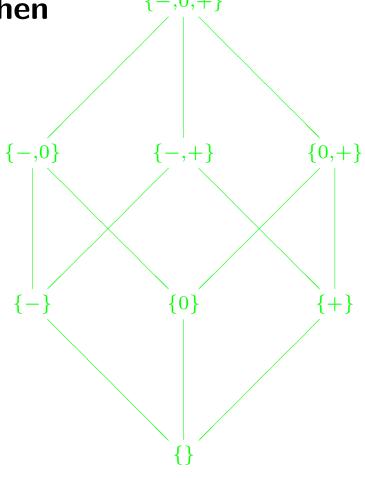
Entsprechend definieren wir die Konkretisierung  $\gamma$  durch

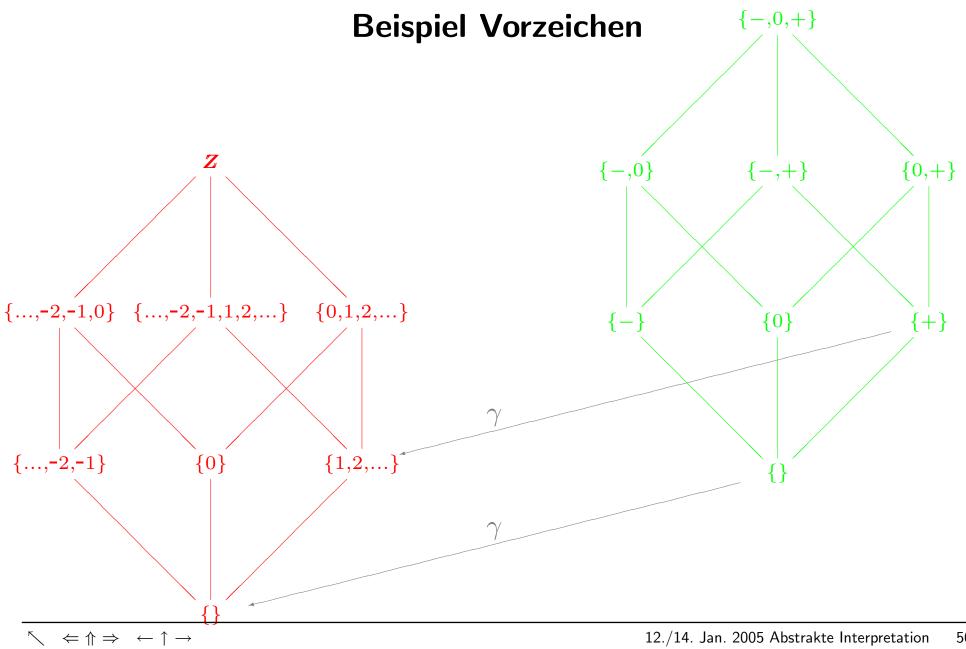
$$\gamma(\{-\}) = \{\dots, -3, -2, -1\}$$
  
 $\gamma(\{0\}) = \{0\}$   
 $\gamma(\{+\}) = \{1, 2, 3, \dots\}$   
 $\gamma(S') = \bigcup_{s \in S'} \gamma(\{s\}).$ 

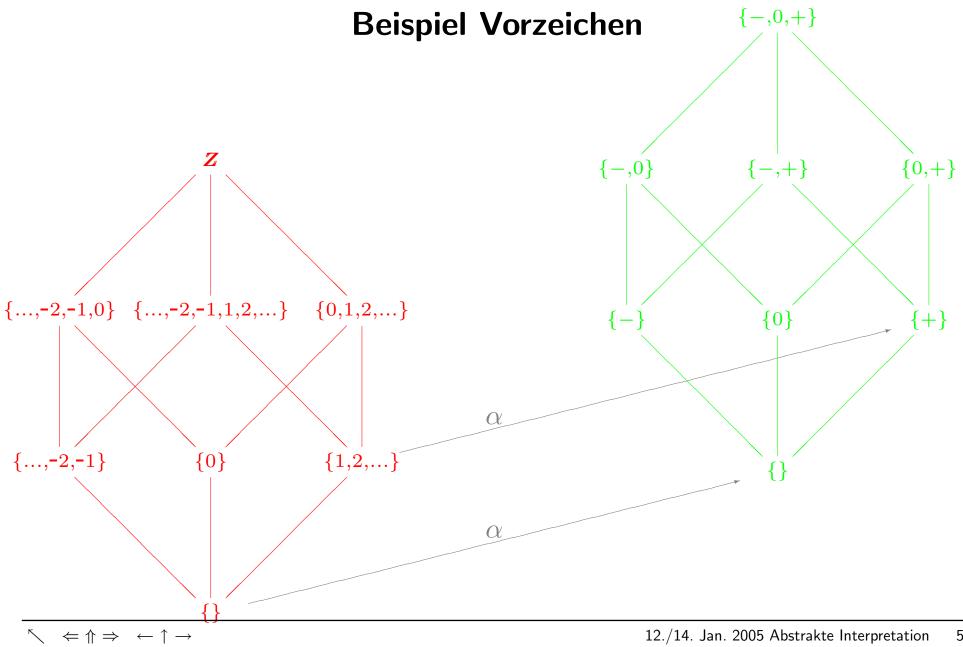
$$\begin{array}{l} \mathbf{Z}, \mathbf{B}, \\ \alpha(\{-2, -1\}) = \alpha(\{-2\}) \cup \alpha(\{-1\}) = \{-\} \cup \{-\} = \{-\}, \\ \alpha(\{0, 2, 4, 6\}) = \gamma(\{0\}) \cup \{0\} \cup \{2\}) \cup \{4\}) \cup \alpha(\{6\}) \cup \ldots = \{0, +\}. \\ = \{\ldots, -3, -2, -1\} \cup \{0\} \cup \{1, 2, 3, \ldots\} = \mathbf{Z}. \end{array}$$

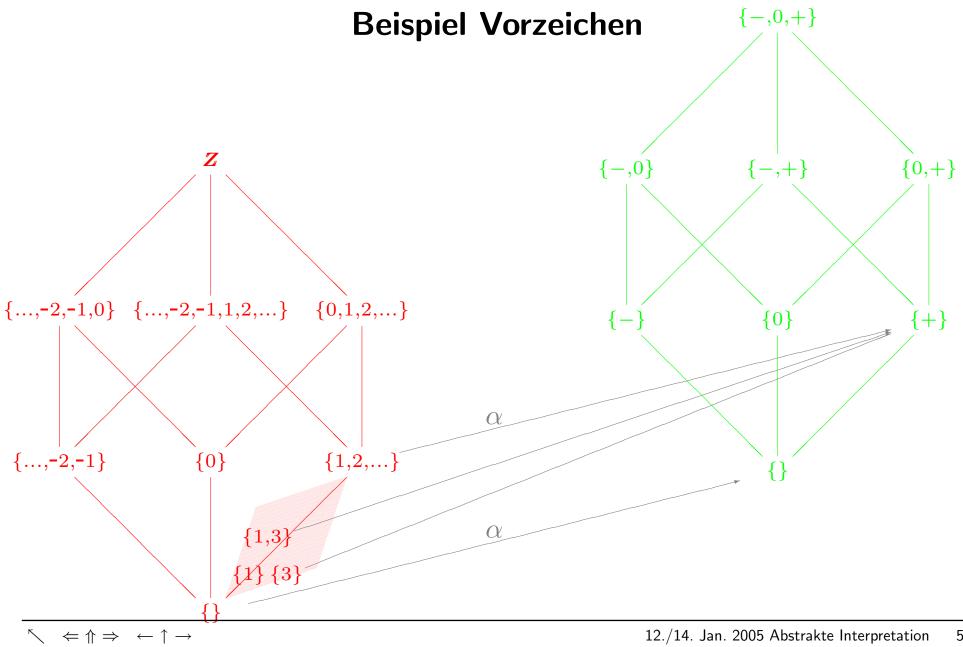


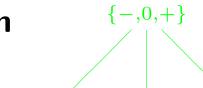


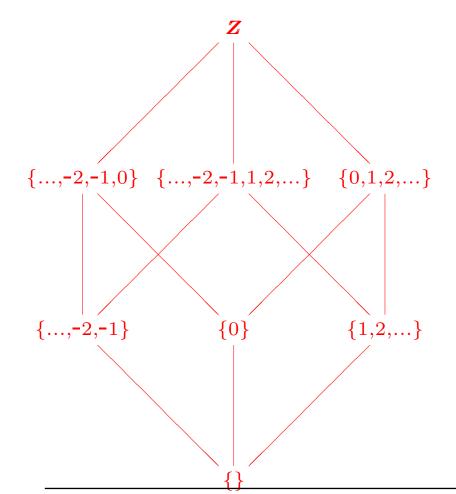


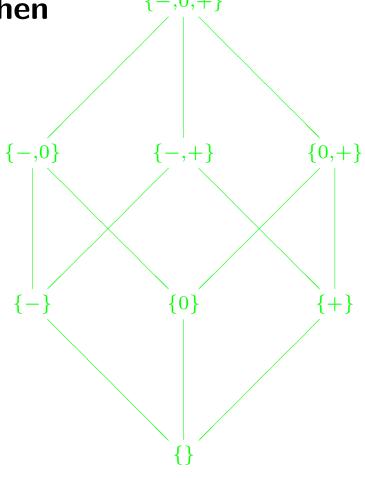


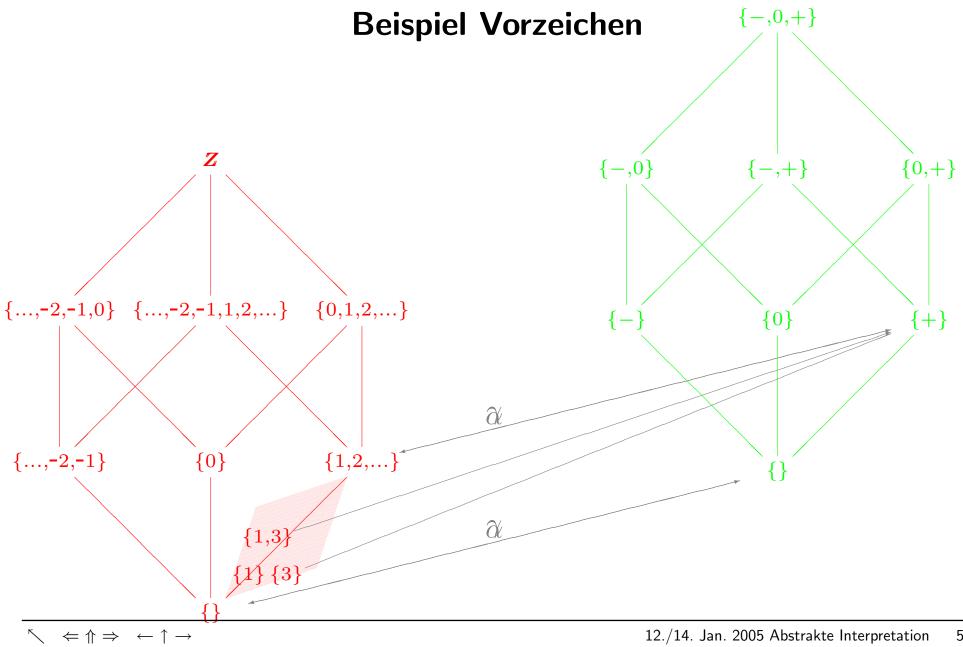












#### Beispiel Restklassen

```
M = \wp(Z) mit (\subseteq), M' = \wp(\{0, \dots, m-1\}) mit (\subseteq) für ein festes m \ge 2. Wir definieren \alpha(\{n\}) = \{n\%m\} für einelementige, \alpha(S) = \bigcup_{n \in S} \alpha(\{n\}) für beliebige Mengen; sowie \gamma(\{n'\}) = \{n \in I\!\!N \mid n\%m = n'\} und \gamma(S') = \bigcup_{s \in S'} \gamma(\{s\}).
```

Wir bilden also alle (konkreten) Zahlen, die denselben Rest mod. m haben, auf dieselbe (abstrakte) Zahl ab.

Für m=2 ist  $0\in M'$  bzw.  $1\in M'$  die Abstraktion der geraden bzw. der ungeraden Zahlen.

Für m=256 ist  $0,\ldots,255\in M'$  jeweils die Abstraktion der Zahlen mit dem entsprechenden Wert des niederwertigsten Byte, z.B.

 $\alpha(\{0x2900, 0x0a00, 0x00fe, 0x12fe, 0x85fe\}) = \{0, 254\}.$ 

## Beipiel Restklassen: Eigenschaften

| Name            | Eigenschaft                                       |                            |
|-----------------|---|----------------------------|
| $\gamma \neq$   | $\gamma(\{a'\}) \cap \gamma(\{b'\}) = \{\}$       | $f\ddot{u} r \ a' \neq b'$ |
| $\gamma \cap$   | $\gamma(x' \cap y') = \gamma(x') \cap \gamma(y')$ |                            |
| $\gamma$ $\cup$ | $\gamma(x' \cup y') = \gamma(x') \cup \gamma(y')$ |                            |
| $\alpha \cup$   | $\alpha(x \cup y) = \alpha(x) \cup \alpha(y)$     |                            |
| $\alpha \gamma$ | $\alpha(\gamma(x')) = x'$                         |                            |

Satz (Cousot, Cousot 1976):

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq),\bot und M' mit (\sqsubseteq'),\bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: \stackrel{\alpha}{M} \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist.
```

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq),\bot und M' mit (\sqsubseteq'),\bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: M \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist. In jedem Fall gilt aber X \sqsubseteq \gamma(X').
```

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq), \bot und M' mit (\sqsubseteq'), \bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: M \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist. In jedem Fall gilt aber X \sqsubseteq \gamma(X').
```

D.h. wir können im abstrakten Verband (mit  $\Phi'$ ) genauso rechnen wie vorher im konkreten (mit  $\Phi$ , "collecting semantics").

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq),\bot und M' mit (\sqsubseteq'),\bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: M \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist. In jedem Fall gilt aber X \sqsubseteq \gamma(X').
```

D.h. wir können im abstrakten Verband (mit  $\Phi'$ ) genauso rechnen wie vorher im konkreten (mit  $\Phi$ , "collecting semantics"). Der abstrakte Punkt X' ist eine obere Approximation des konkreten Fixpunkts X.

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq), \bot und M' mit (\sqsubseteq'), \bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: M \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist. In jedem Fall gilt aber X \sqsubseteq \gamma(X').
```

D.h. wir können im abstrakten Verband (mit  $\Phi'$ ) genauso rechnen wie vorher im konkreten (mit  $\Phi$ , "collecting semantics"). Der abstrakte Punkt X' ist eine obere Approximation des konkreten Fixpunkts X. Wir haben ein Rezept bekommen für die Überführung der Programmkonstrukte.

```
Satz (Cousot, Cousot 1976): Seien M mit (\sqsubseteq),\bot und M' mit (\sqsubseteq'),\bot' vollständige Verbände. Sei M \stackrel{\gamma}{\Longrightarrow} M' eine Galois-Verbindung. Sei \Phi: M \longrightarrow M monoton und stetig. Sei \Phi': M' \longrightarrow M' monoton, so daß \alpha(\Phi(\gamma(x'))) \sqsubseteq' \Phi'(x'). Nach dem Fixpunktsatz ist dann X = \bigsqcup \{\bot, \Phi(\bot), \Phi(\Phi(\bot)), \ldots\} der kleinste Fixpunkt von \Phi. X' = \bigsqcup' \{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\} ist nur dann der kleinste Fixpunkt von \Phi', wenn \Phi' zusätzlich stetig ist. In jedem Fall gilt aber X \sqsubseteq \gamma(X').
```

D.h. wir können im abstrakten Verband (mit  $\Phi'$ ) genauso rechnen wie vorher im konkreten (mit  $\Phi$ , "collecting semantics"). Der abstrakte Punkt X' ist eine obere Approximation des konkreten Fixpunkts X. Wir haben ein Rezept bekommen für die Überführung der Programmkonstrukte.

## Anforderungen: Überführbarkeit

Jedes Programmkonstrukt (Verzweigung, Zusammenführung, Zuweisung, . . . ) muß sich in eine abstrakte Operation / Gleichung überführen lassen. Insbesondere die im Programm auftretenden Operationen müssen eine abstrakte Entsprechung haben.

## Überführung von Programmkonstrukten am Beispiel

Bei der Berechnung der collecting semantics in unserem Beispiel hatten wir als (konkreten) Verband  $M = \wp(\mathbf{Z})^8$  gewählt, entsprechend den 8 Programmpunkten  $A, \ldots, H$ , an denen wir die Wertemengen der int-Variablen x bestimmen wollten.

Jetzt wollen wir z.B. nur wissen, welche Restklassen mod. 6 an einem Programmpunkt auftreten können

(daraus können wir die Reste mod. 2 und mod. 3 ablesen; 2 und 3 treten als einzige Konstanten im Programmtext auf).

Wir wählen daher als abstrakten Verband  $M' = \wp(\{0, \dots, 5\})^8$ , dazu  $\alpha$  und  $\gamma$  komponentenweise wie auf der obigen Folie.

# Überführung von Programmkonstrukten am Beispiel

Damit können den Mindestwert von  $\Phi'$  einfach ausrechnen:

```
\begin{array}{ll} &\alpha(\Phi(\gamma(\langle A',\ldots,H'\rangle)))\\ =&\alpha(\Phi(\langle \gamma(A'),\ldots,\gamma(H')\rangle)) &\gamma \text{ komponentenweise}\\ =&\alpha(\langle\{1,\ldots,5\},\ldots,\{3n+1\mid n\in\gamma(F')\}\rangle) &\text{Def. }\Phi\\ =&\langle\alpha(\{1,\ldots,5\}),\\ &\alpha(\gamma(A')\cup\gamma(G')\cup\gamma(H')),\\ &\alpha(\gamma(B')\cap\{2,3,4,\ldots\}),\\ &\alpha(\gamma(B')\cap\{\ldots,-2,-1,0,1\}),\\ &\alpha(\gamma(C')\cap\{\ldots,-4,-2,0,2,4,\ldots\}),\\ &\alpha(\gamma(C')\cap\{\ldots,-3,-1,1,3,\ldots\}),\\ &\alpha(\{n/2\mid n\in\gamma(E')\}),\\ &\alpha(\{3n+1\mid n\in\gamma(F')\})\rangle &\alpha \text{ komponentenweise} \end{array}
```

# Überführung von Programmkonstrukten am Beispiel

Jetzt schätzen wir die einzelnen Komponenten möglichst knapp nach oben ab:

$$\begin{array}{lll} \alpha(\{1,...,5\}) & = \{1,...,5\} & \text{Def. } \alpha \\ \alpha(\gamma(A') \cup \gamma(G') \cup \gamma(H')) & = \alpha(\gamma(A' \cup G' \cup H')) & \gamma \cup \\ & = A' \cup G' \cup H' & \alpha \gamma \\ \alpha(\gamma(B') \cap \{2,3,4,...\}) & \sqsubseteq' \alpha(\gamma(B')) \\ & = B' & \alpha \gamma \\ \alpha(\gamma(B') \cap \{...,-2,-1,0,1\}) & \sqsubseteq' \alpha(\gamma(B')) \\ & = B' & \alpha \gamma \\ \alpha(\gamma(C') \cap \{...,-4,-2,0,2,4,...\}) & = \alpha(\gamma(C' \cap \{0,2,4\})) & \gamma \cap \\ & = C' \cap \{0,2,4\} & \alpha \gamma \\ \alpha(\gamma(C') \cap \{...,-3,-1,1,3,...\}) & = \alpha(\gamma(C' \cap \{1,3,5\})) & \gamma \cap \\ & = C' \cap \{1,3,5\} & \alpha \gamma \end{array}$$

# Überführung von Programmkonstrukten am Beispiel

$$\begin{array}{lll} \alpha(\{n/2 \mid n \in \gamma(E')\}) & \sqsubseteq' \alpha(\mathbf{Z}) \\ & = \{0,...,5\} \\ \alpha(\{3n+1 \mid n \in \gamma(F')\}) = \alpha(\{3n+1 \mid n' \in F' \land n \in \gamma(\{n'\})\}) & \gamma \cup \\ & = \bigcup_{n' \in F'} \bigcup_{n \in \gamma(\{n'\})} \alpha(\{3n+1\}) & \alpha \cup \\ & = \bigcup_{n' \in F'} \bigcup_{n \in \gamma(\{n'\})} \{(3n+1)\%6\} & \alpha \ \{\cdot\} \\ & = \bigcup_{n' \in F'} \bigcup_{n \in \gamma(\{n'\})} \{(3(n\%6)+1)\%6\} & \text{Arithmetik} \\ & = \bigcup_{n' \in F'} \bigcup_{n \in \gamma(\{n'\})} \{(3n'+1)\%6\} & \gamma \ \{\cdot\} \\ & = \bigcup_{n' \in F'} \{(3n'+1)\%6\} & \text{alle gleich} \\ & = \{(3n'+1)\%6 \mid n' \in F'\} \end{array}$$

# Überführung von Programmkonstrukten am Beispiel

Wir können also definieren:

$$\Phi'(A', \dots, H') = \langle \{1, \dots, 5\}, \\ A' \cup G' \cup H', \\ B', \\ B', \\ C' \cap \{0, 2, 4\}, \\ C' \cap \{1, 3, 5\}, \\ \{0, \dots, 5\}, \\ \{(3n'+1)\%6 \mid n' \in F'\} \rangle$$

Dabei müssen wir darauf achten, daß  $\Phi'$  monoton wird (das ist hier der Fall).

Diese Definition von  $\Phi'$  enthält nur noch abstrakte Konstrukte.

# Anforderungen: Terminierung

nochmal Satz (Cousot, Cousot 1976):

Der kleinste Fixpunkt X der collecting semantics kann abgeschätzt werden durch  $X \sqsubseteq \gamma(X')$ , wobei  $X' = | |'\{\bot', \Phi'(\bot'), \Phi'(\Phi'(\bot')), \ldots\}$ .

Wir wollen also sicherstellen, daß die Berechnung von X' immer terminiert. Dazu gibt es verschiedene Möglichkeiten:

- 1. Der abstrakte Verband M' hat nur endlich viele Elemente.
- 2. Jede unendlich lange aufsteigende Kette in M' wird irgendwann stationär ("Ascending Chain Condition (ACC)", s.u.).
- 3. Wir können  $\Phi'$  so konstruieren, daß die Terminierung gesichert ist ("Widening", s.u.).

### **Terminierung: Ascending Chain Condition**

ACC: jede unendlich lange aufsteigende Kette in M' wird irgendwann stationär.

#### Formal:

für jede aufsteigende Kette  $x_1' \sqsubseteq' x_2' \sqsubseteq' \ldots \sqsubseteq' x_i' \sqsubseteq' \ldots$  existiert ein  $n \in I\!\!N$ , so daß  $x_n' = x_{n+1}' = \ldots = x_{n+j}' = \ldots$ 

Wir nennen die Kette in diesem Fall stationär ab n.

In diesem Fall ist  $\bigsqcup'\{x'_1, x'_2, \ldots, x'_i, \ldots\} = x'_n$ .

Da die Iterationen von  $\Phi'$  eine aufsteigende Kette bilden (wegen der Monotonie von  $\Phi'$  ist  $\bot' \sqsubseteq' \Phi'(\bot') \sqsubseteq' \Phi'(\Phi'(\bot')) \sqsubseteq' \ldots$ ), garantiert ACC, daß sie irgendwann stationär wird und daher nach endlich vielen Iterationen der Maximalwert X' erreicht ist.

### **Ascending Chain Condition: Beispiele**

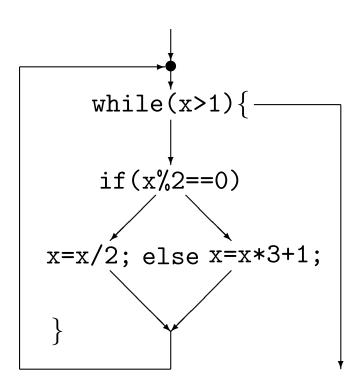
```
 \{-\infty,\ldots,-4,-3,-2,-1,0\} \text{ mit der \"{u}blichen Ordnung } (\leq).   \{\{\},I\!\!N\} \cup \\ \{\{0\},\{1\},\{2\},\ldots\} \cup \\ \{\{0,1\},\{0,2\},\{0,3\},\ldots,\{1,2\},\{1,3\},\ldots,\{2,3\},\ldots\} \\ \text{mit } (\subseteq) \text{ ("ostfriesische Mengen")}
```

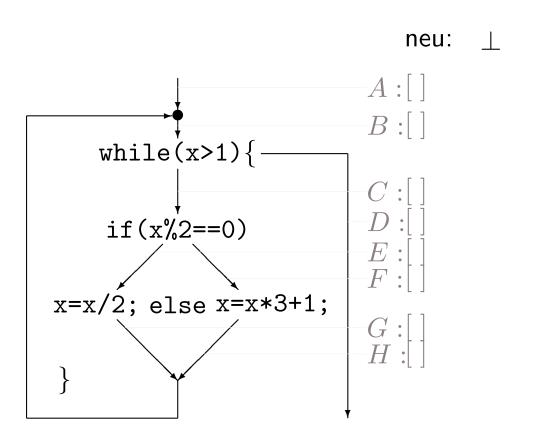
## **Terminierung: Widening**

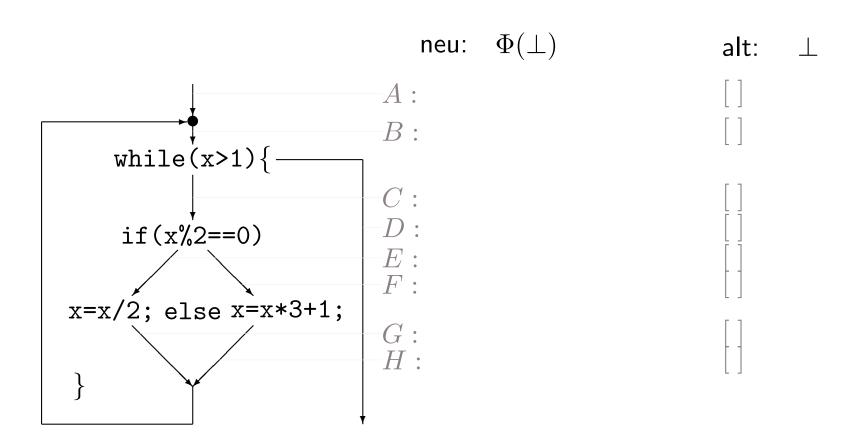
Der Verband der Intervalle ist aber weder endlich, noch erfüllt er ACC.

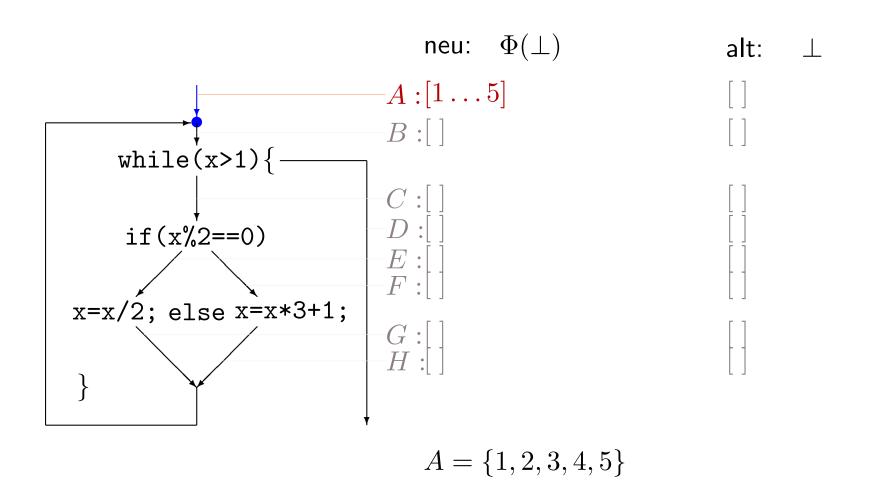
(z.B. 
$$[1,2] \sqsubseteq' [1,3] \sqsubseteq' \ldots \sqsubseteq' [1,i] \sqsubseteq' \ldots$$
).

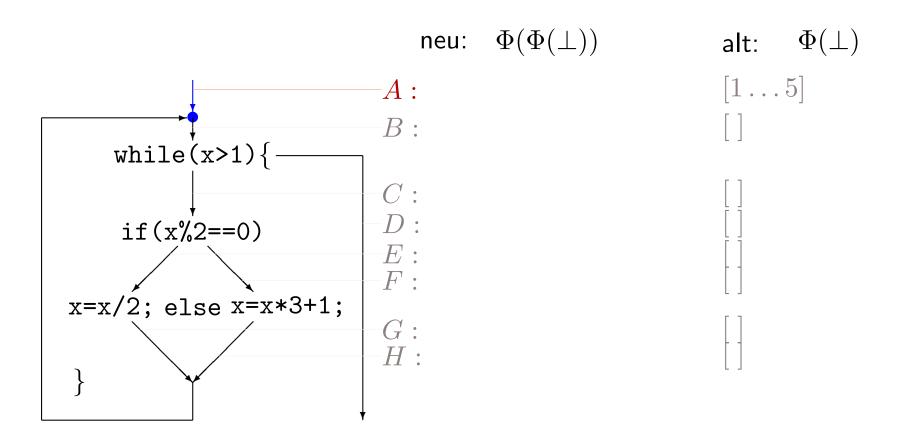
In unserem Beispiel terminiert die Berechnung der abstrakten Interpretation tatsächlich nicht (s.u.).

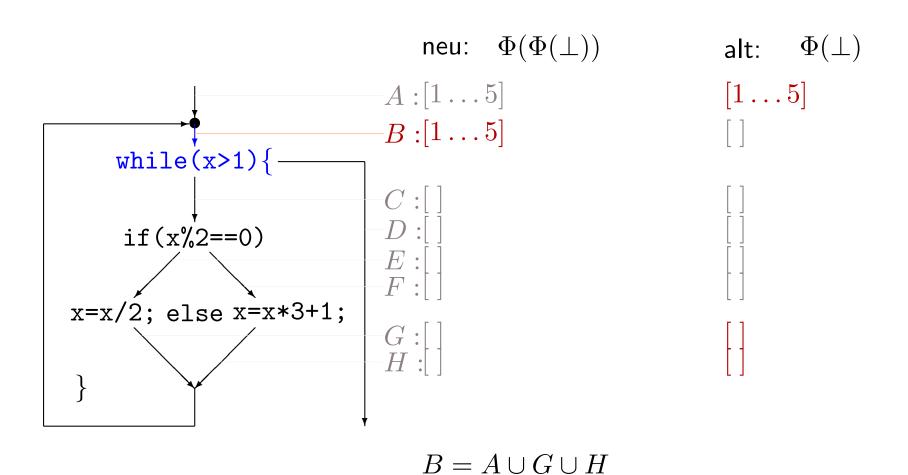


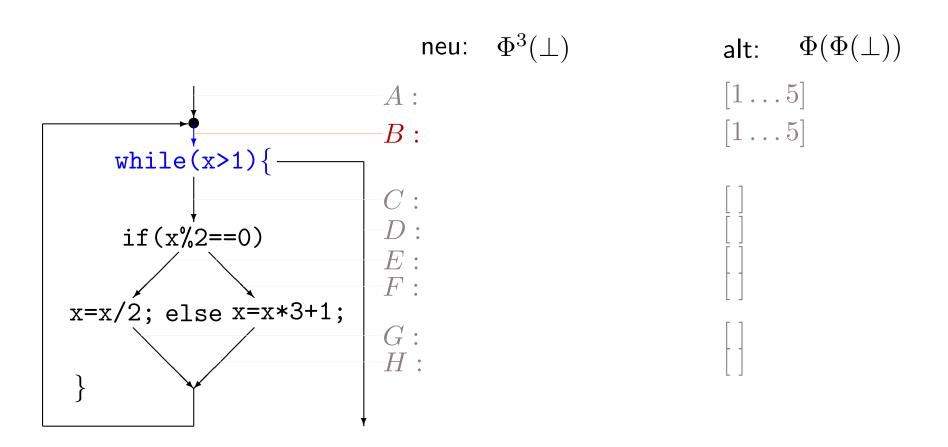


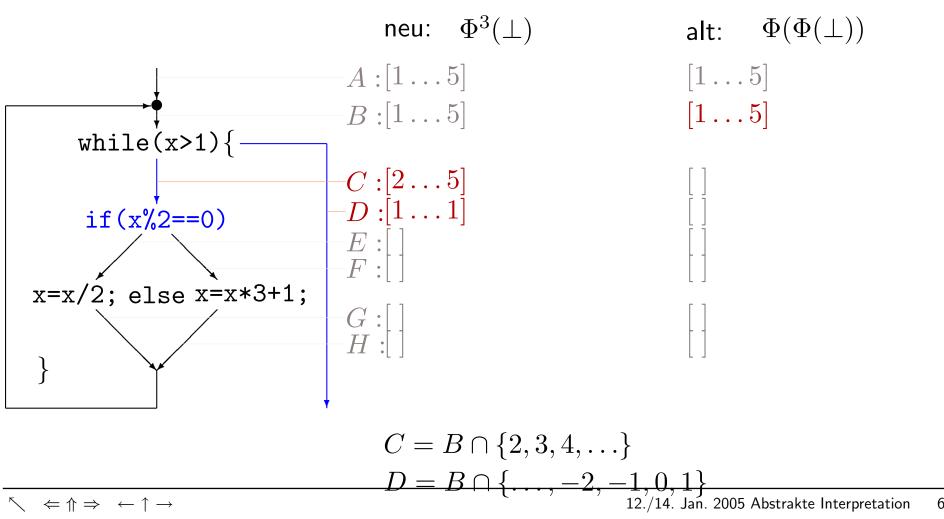


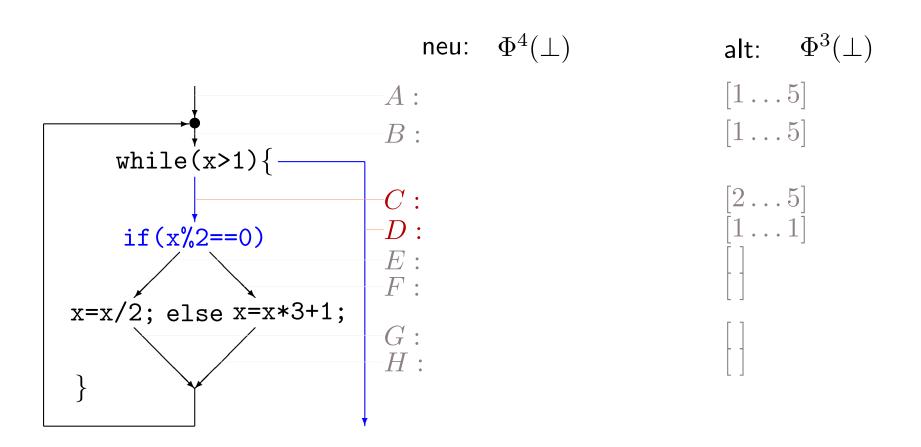


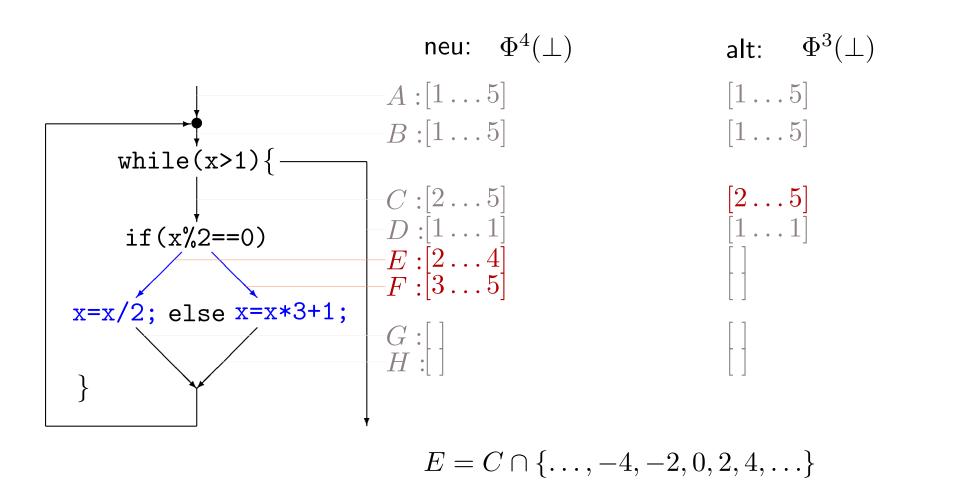


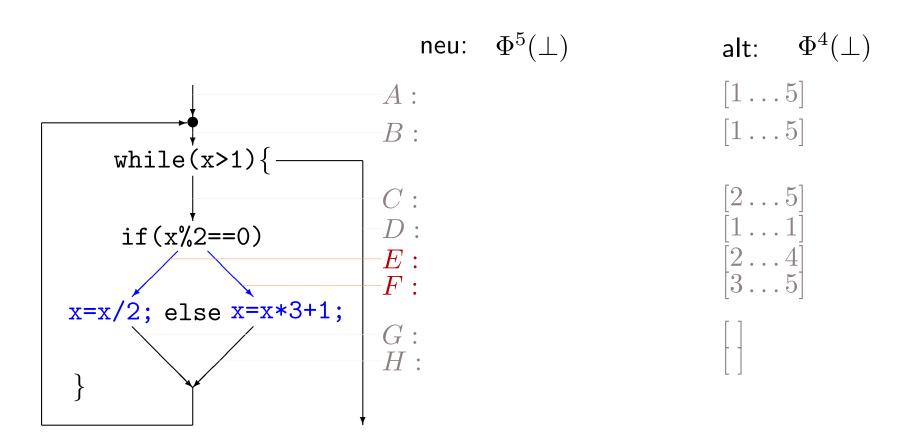


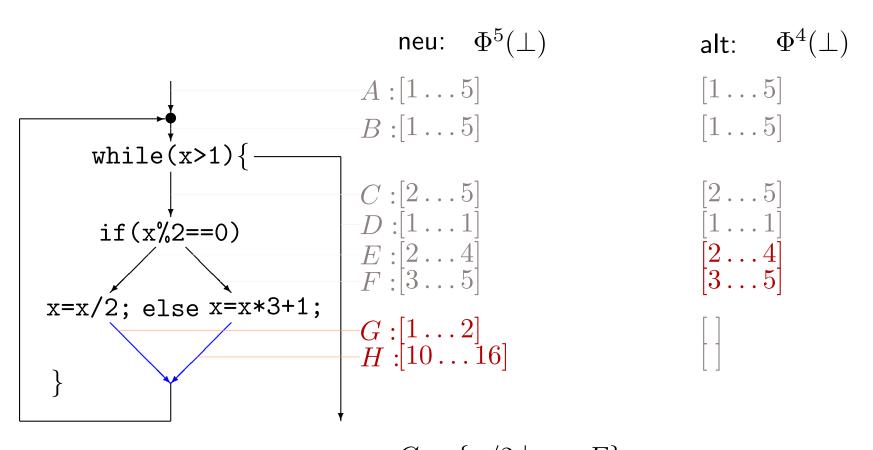






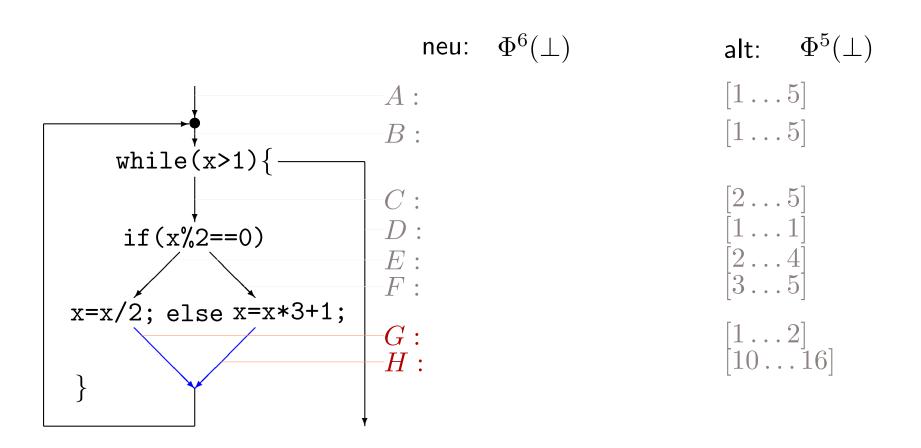


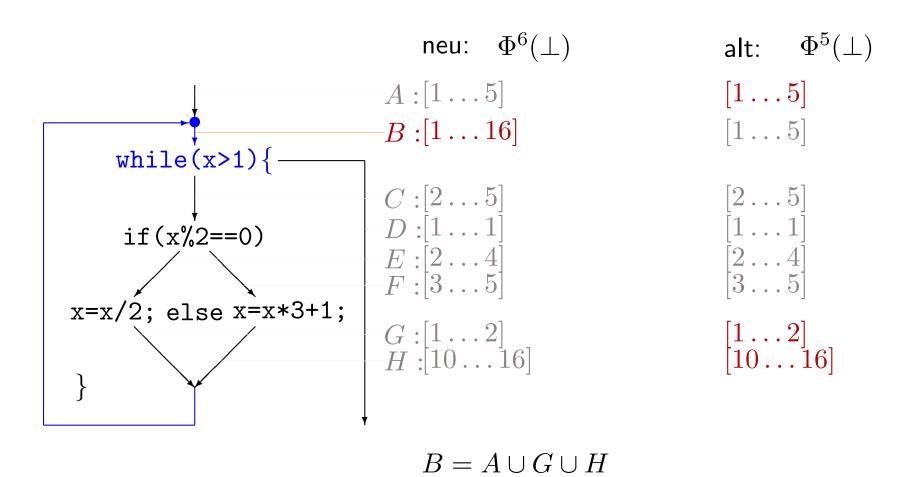


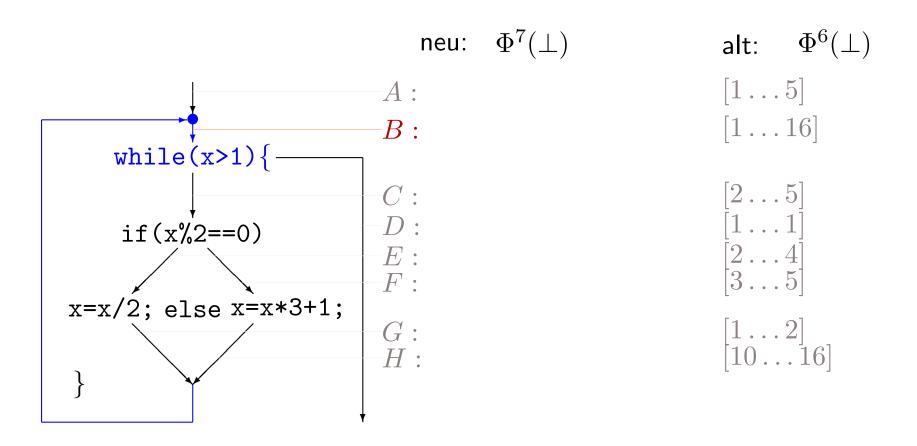


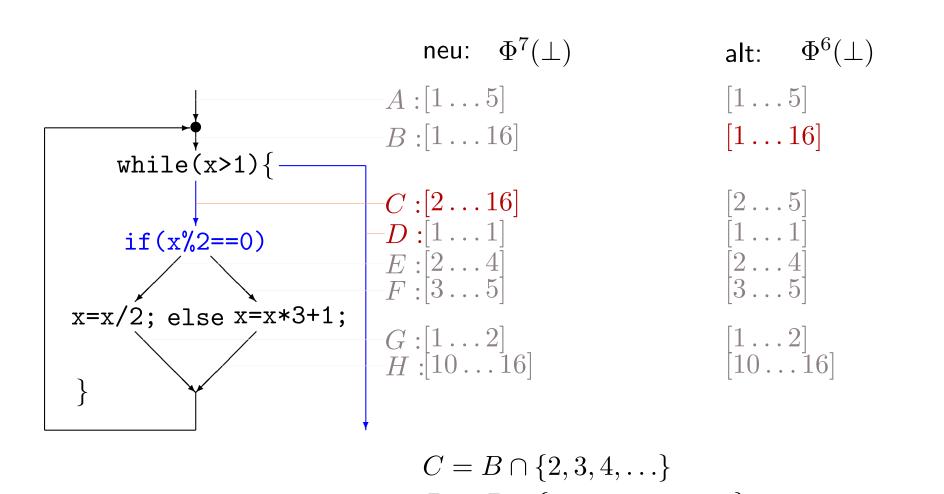
$$G = \{n/2 \mid n \in E\}$$
 
$$H = \{3 \cdot n + 1 \mid n \in F\}$$
 12./14. Jan. 2005 Abstrakte Interpretation

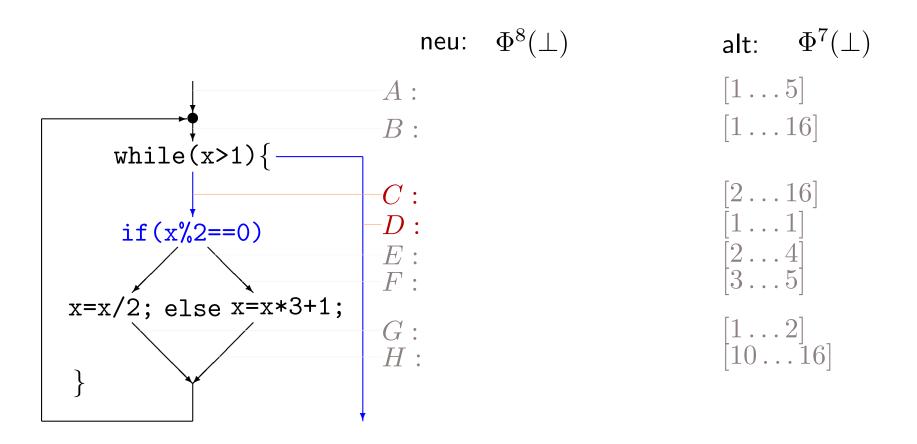
64

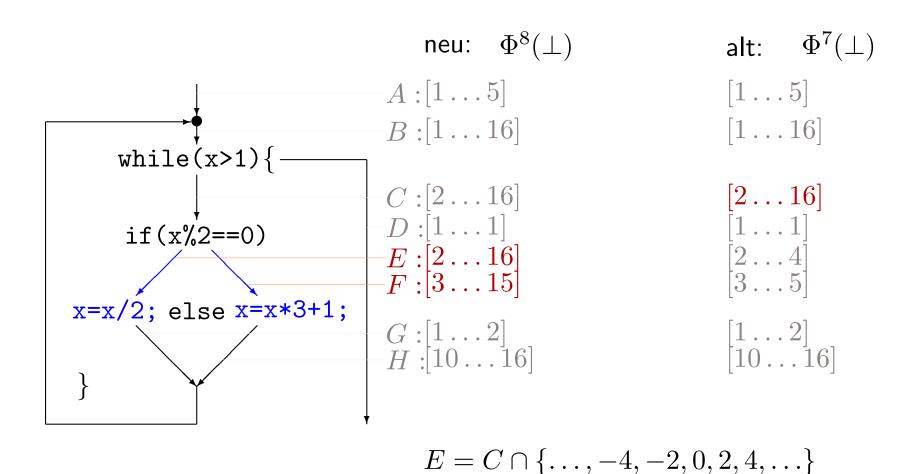


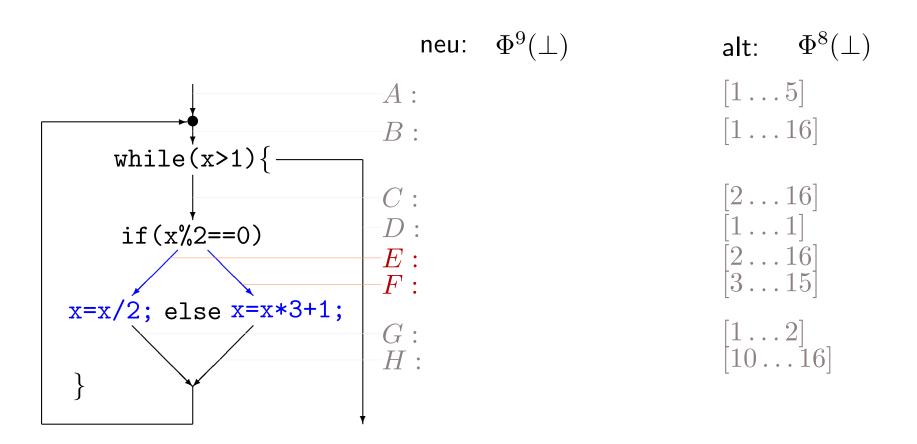


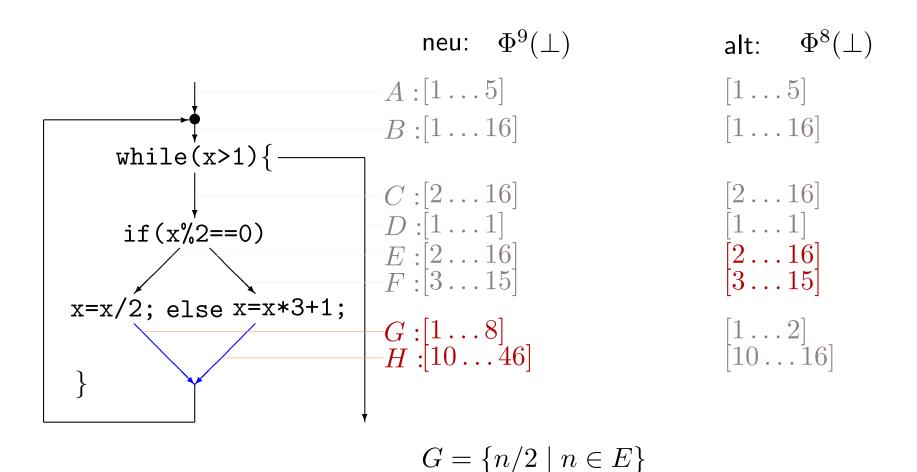


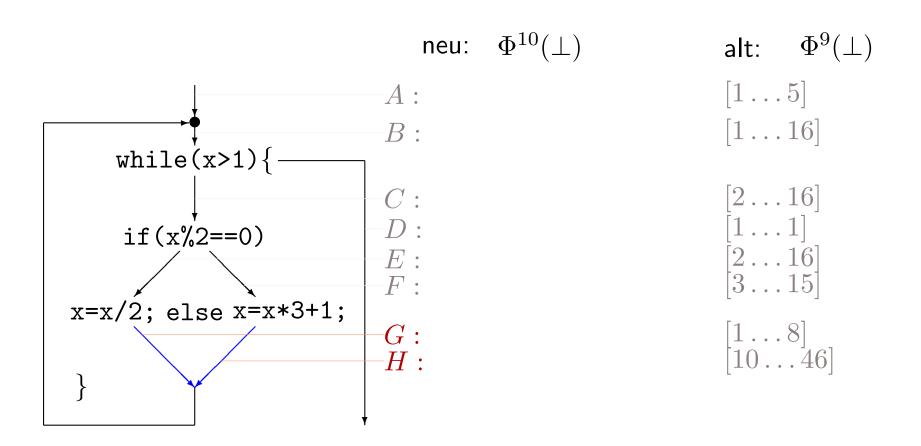


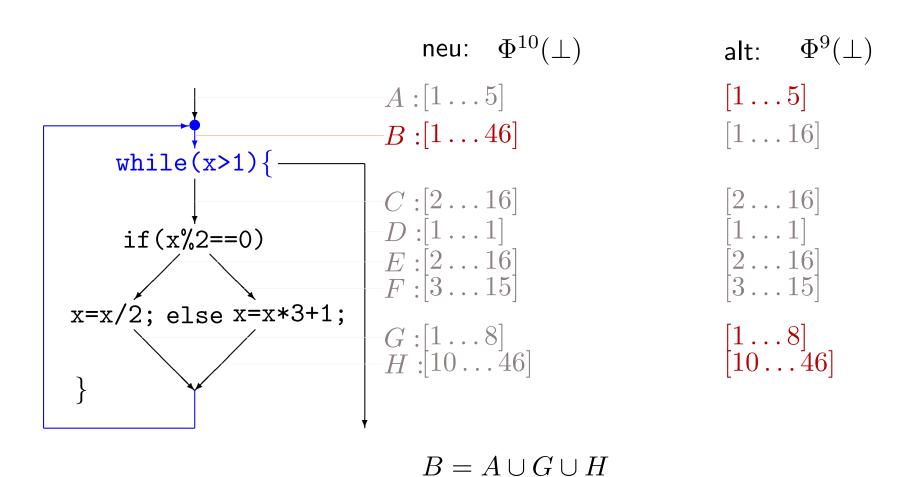


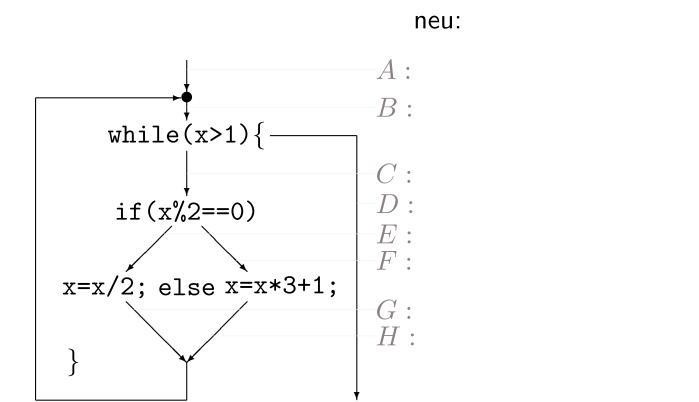




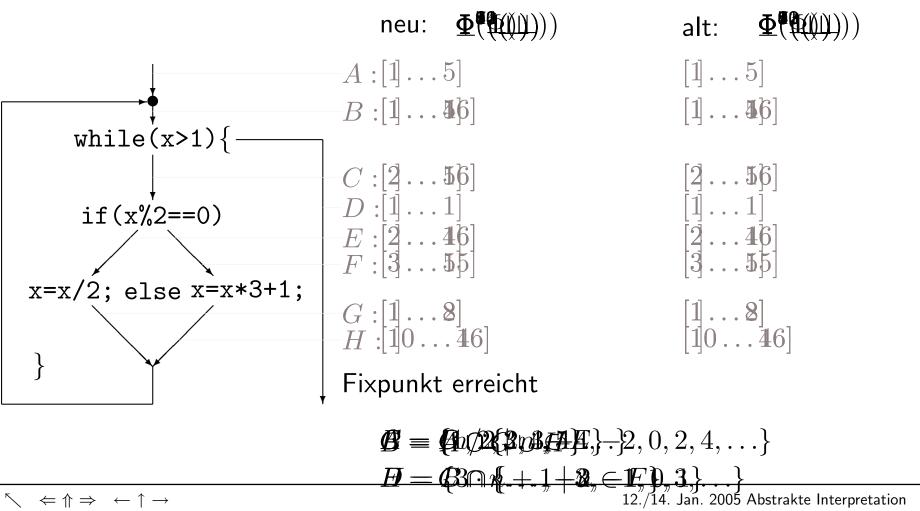








alt:



64

## **Terminierung: Widening**

Andererseits ist es sehr wünschenswert, Informationen über die Grenzen, innerhalb derer sich ein Variablenwert an einer Programmstelle bewegen kann, zu bekommen.

Sofern  $\Phi'$  zusätzlich stetig ist, kann man eine Methode aus (Cousot, Cousot 1992) anwenden, die  $\Phi'$  so modifiziert, daß es nach endlich vielen Iterationen nichts mehr ändert, daß also die Kette  $\bot' \sqsubseteq' \Phi'(\bot') \sqsubseteq' \Phi'(\Phi'(\bot')) \sqsubseteq' \ldots \sqsubseteq' \Phi'^i(\bot') \sqsubseteq' \ldots$  garantiert stationär wird.

## Terminierung: Widening

Formal: Sei  $\Phi'$  zusätzlich stetig.

Sei  $\triangledown: M' \times M' \longrightarrow M'$  eine Operation, so daß  $x' \sqcup y' \sqsubseteq' x' \triangledown y'$  und für jede aufsteigende Kette  $x_1' \sqsubseteq' x_2' \sqsubseteq' \ldots \sqsubseteq' x_i' \sqsubseteq' \ldots$  die Kette  $x_1' \sqsubseteq' x_1' \triangledown x_2' \sqsubseteq' \ldots \sqsubseteq' x_1' \triangledown \ldots \triangledown x_i' \sqsubseteq' \ldots$  irgendwann stationär wird.

Definiere 
$$\Phi''(x') = \begin{cases} x' & \text{für } \Phi'(x') = x' \\ x' \nabla \Phi'(x') & \text{sonst} \end{cases}$$
.

Dann ist die Kette  $\bot' \sqsubseteq' \Phi''(\bot') \sqsubseteq' \Phi''(\Phi''(\bot')) \sqsubseteq' \ldots \sqsubseteq' \Phi''^i(\bot') \sqsubseteq' \ldots$  irgendwann stationär und ihre kleinste obere Schranke  $X'' = \bigsqcup' \{ \bot', \Phi''(\bot'), \Phi''(\Phi''(\bot')), \ldots \}$  ist eine obere Abschätzung für den kleinsten Fixpunkt X' von  $\Phi'$ .

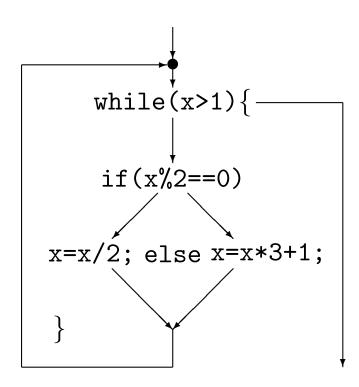
Definiere eine widening-Operation auf Intervallen durch

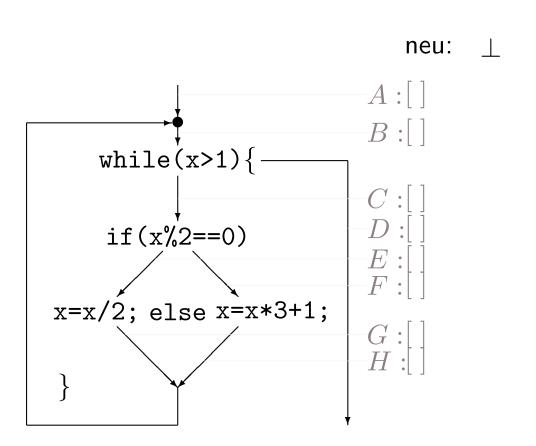
$$[l \dots u] \, \triangledown[\,\,] = [\,\,] \, \triangledown[l \dots u] = [l \dots u] \text{ sowie } [l_1 \dots u_1] \, \triangledown[l_2 \dots u_2] = [l \dots u] \text{ mit } \\ l = \left\{ \begin{array}{ll} -\infty & \text{für} & l_2 < l_1 \\ l_1 & \text{sonst} \end{array} \right. \quad \text{und} \quad u = \left\{ \begin{array}{ll} +\infty & \text{für} & u_2 > u_1 \\ u_1 & \text{sonst} \end{array} \right.$$

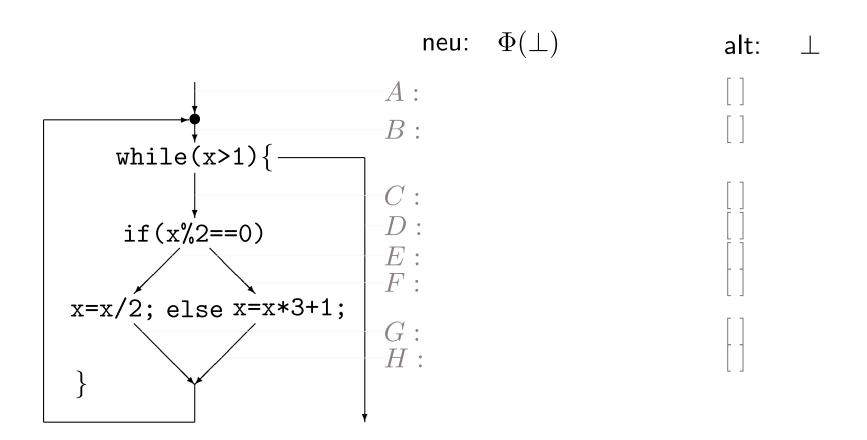
Anschaulich: Sobald sich eine Intervallgrenze verschlechtert, wird sie gleich auf den schlechtestmöglichen Wert gesetzt.

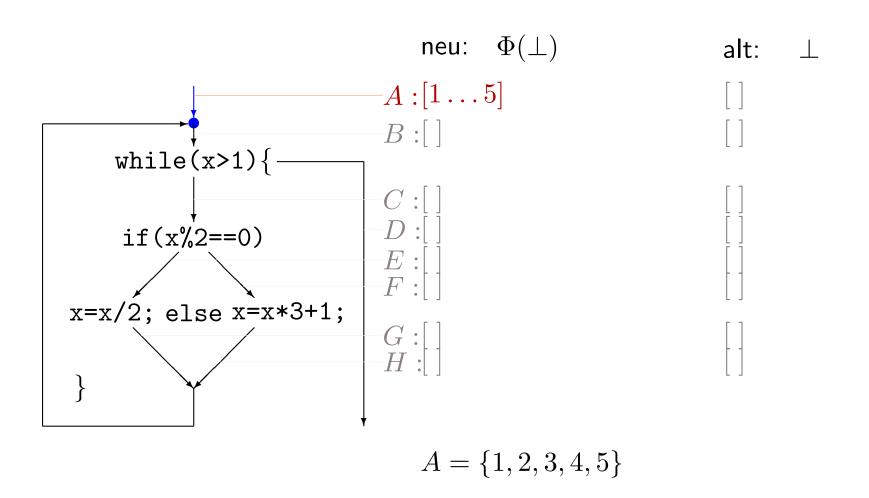
Z.B. 
$$[1...5] \nabla [1...16] = [1...\infty], [1...16] \nabla [1...5] = [1...16].$$

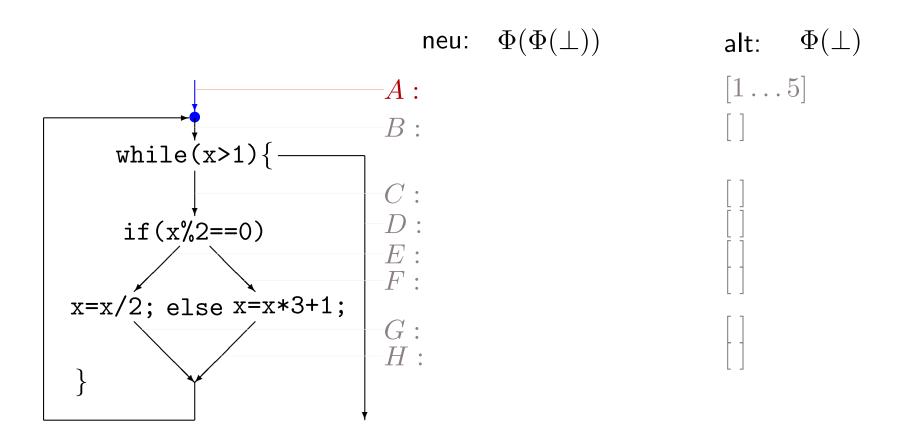
 $(\nabla)$  ist nicht kommutativ, da es die zeitliche Reihenfolge berücksichtigen soll.

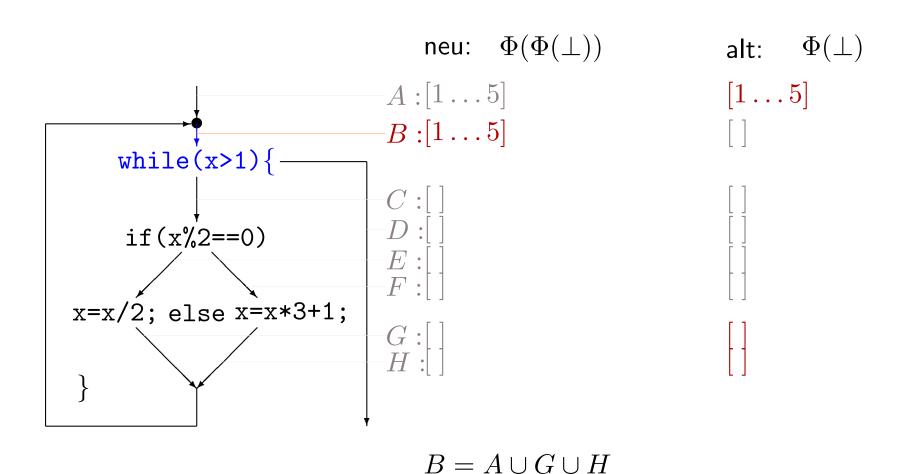


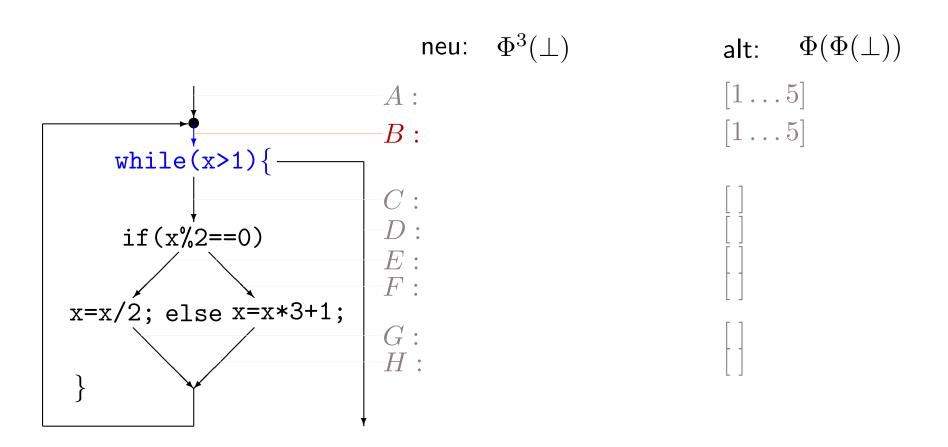


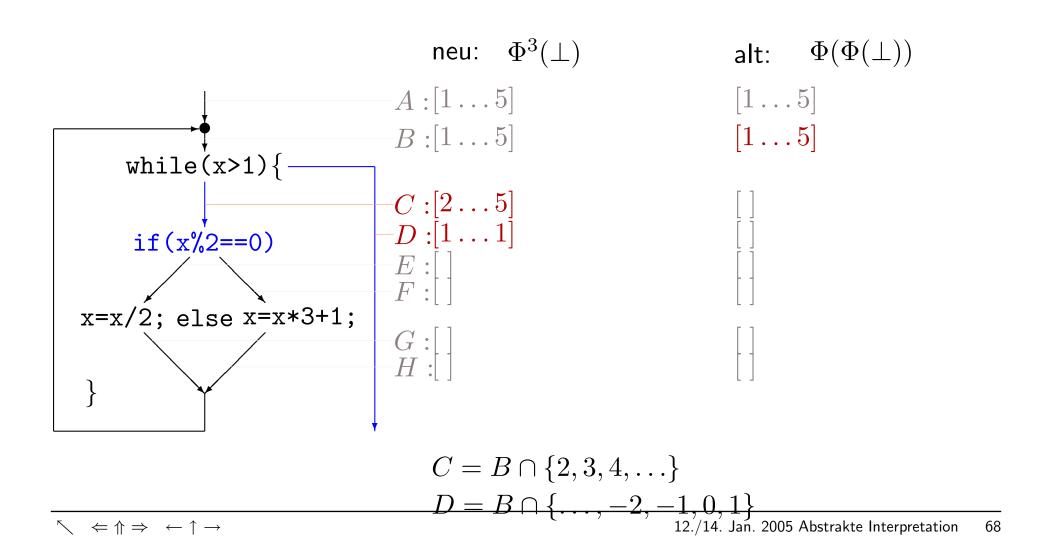


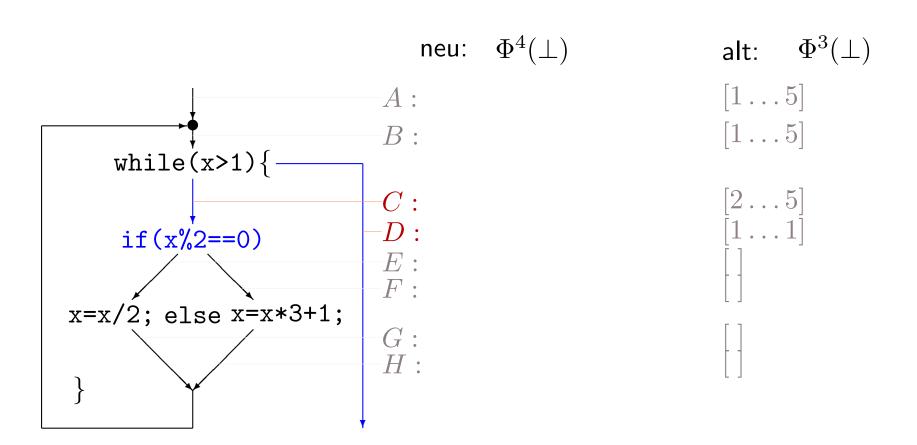


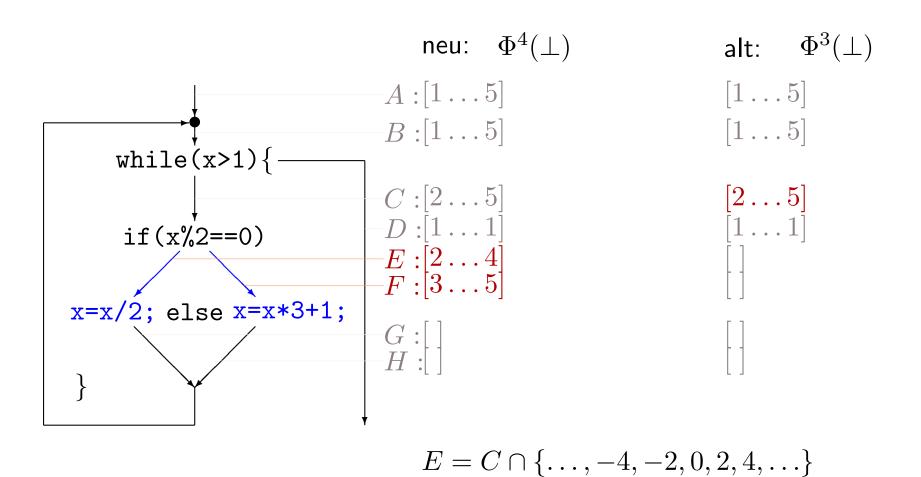




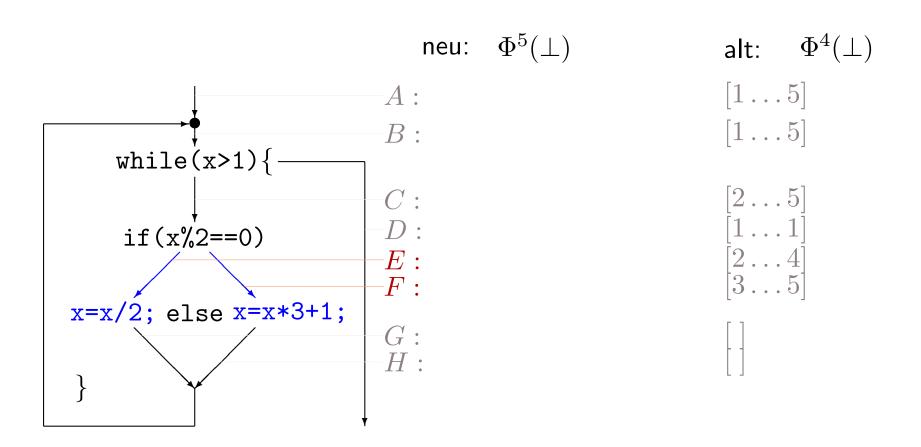


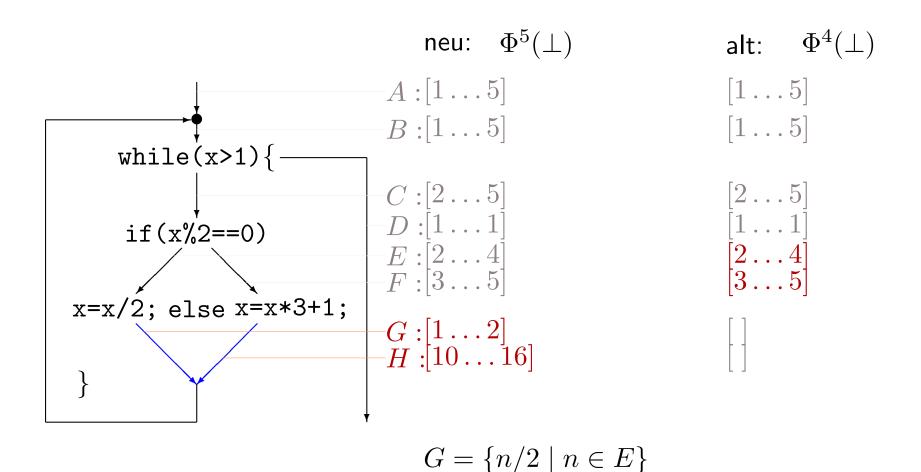




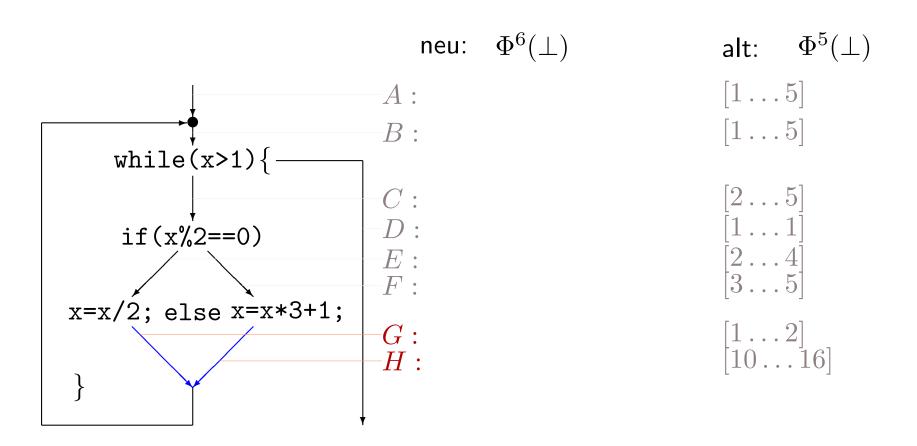


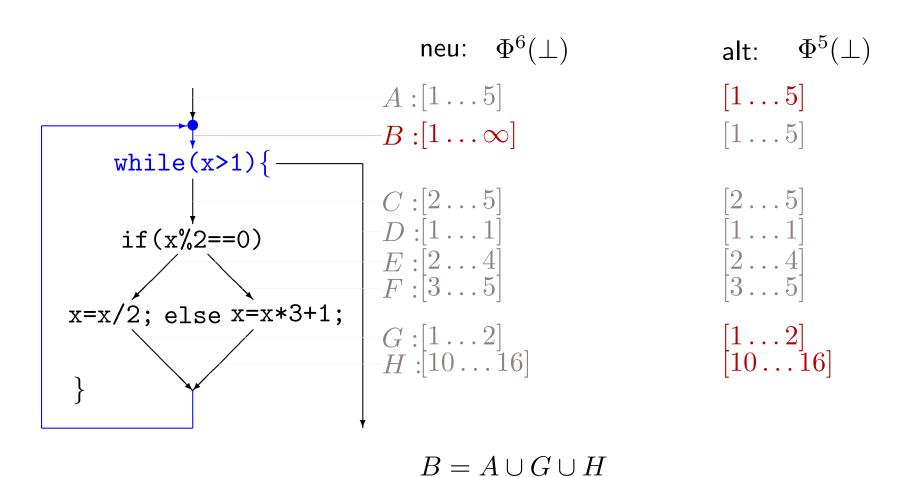
$$^{\nwarrow}$$
  $\Leftarrow \uparrow \Rightarrow \leftarrow \uparrow \rightarrow$ 

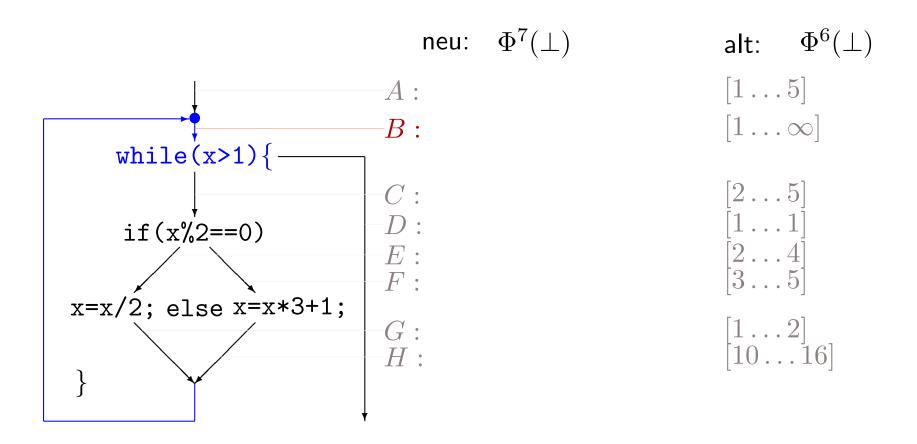


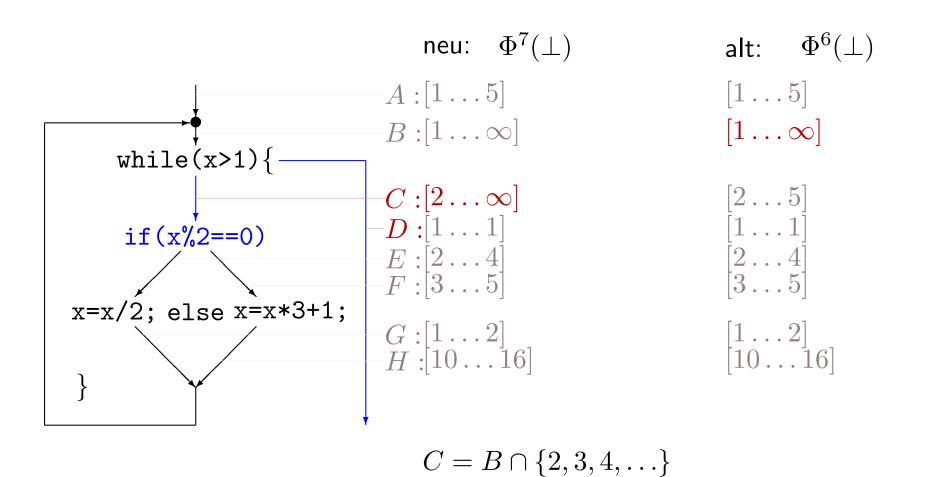


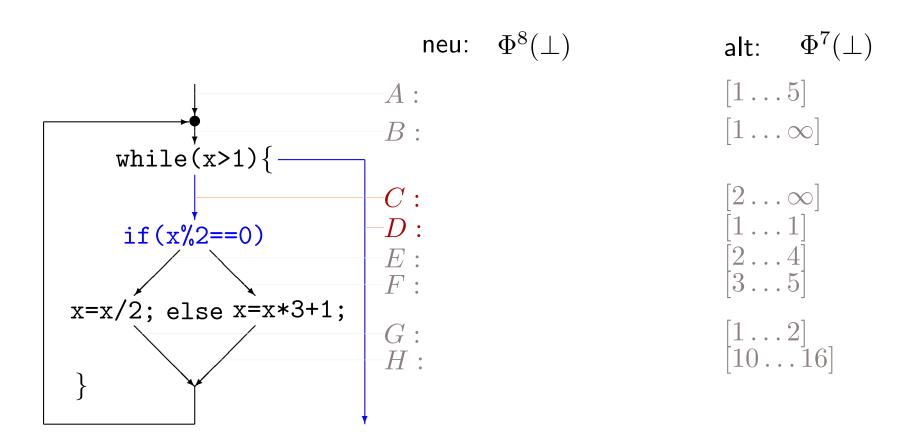
68

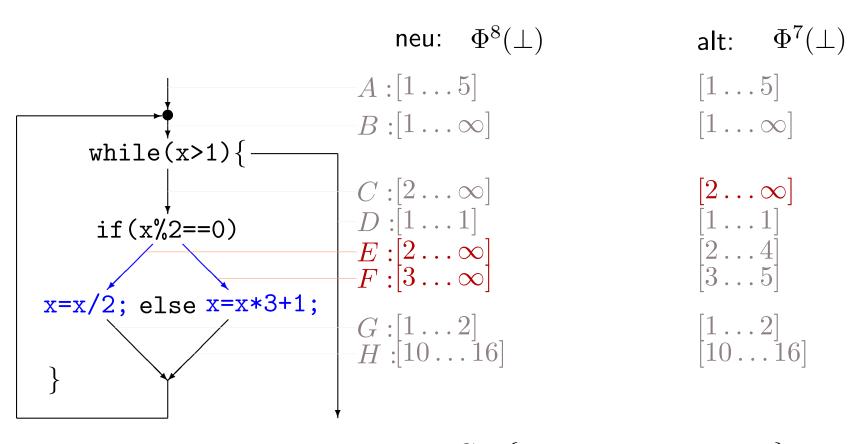






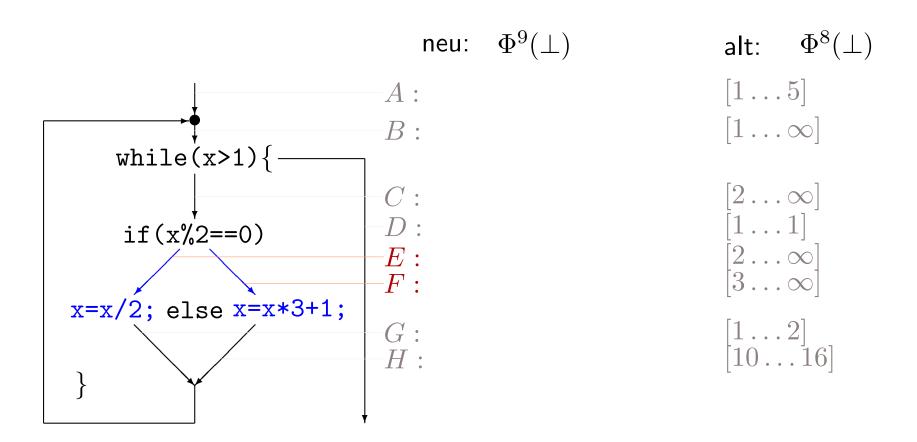


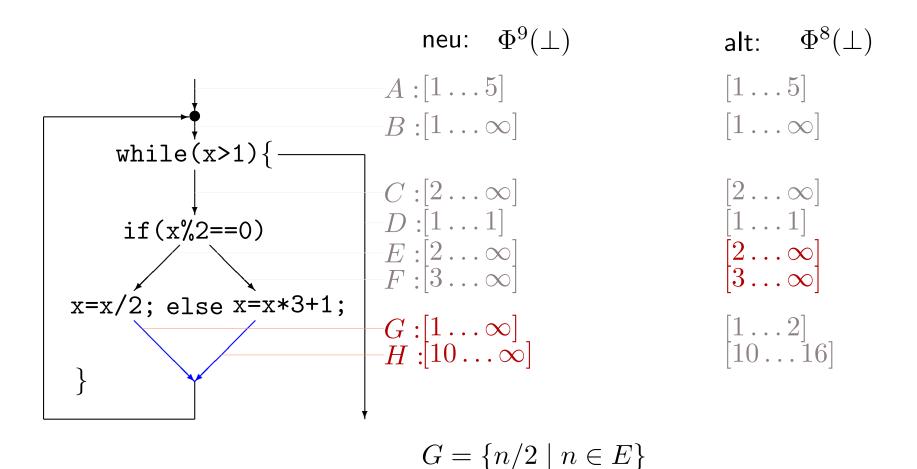




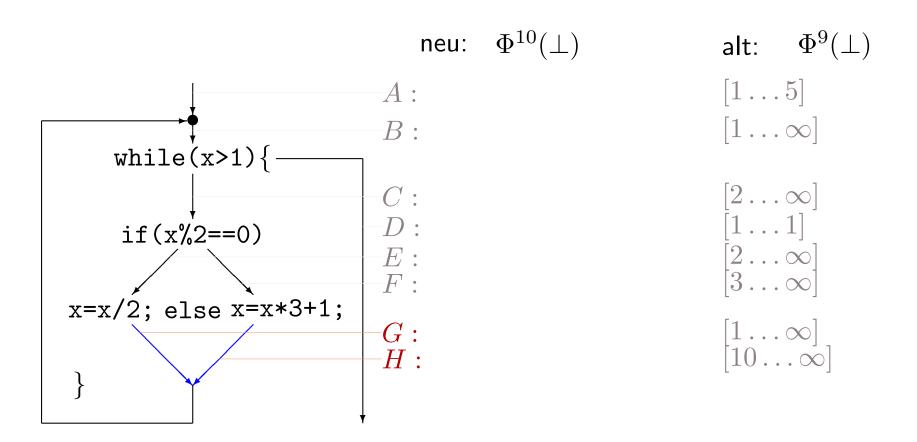
$$E = C \cap \{\dots, -4, -2, 0, 2, 4, \dots\}$$

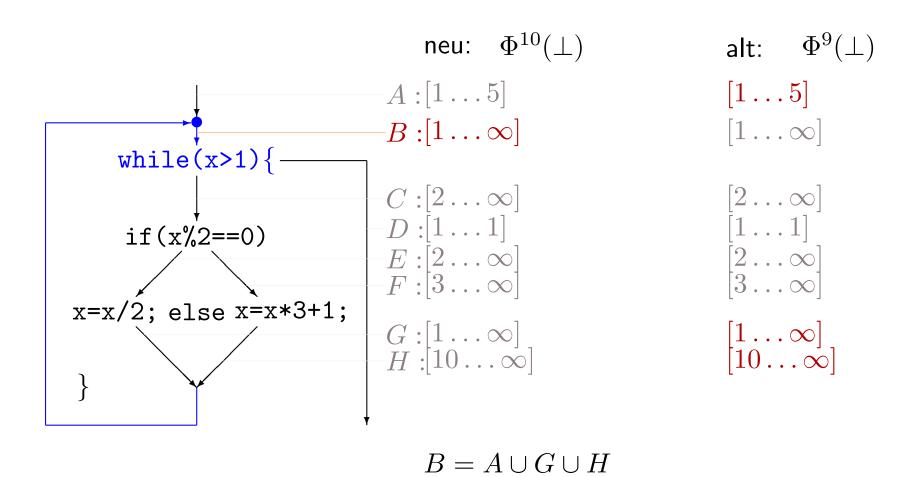
$$F = C \cap \{\dots, -3, -1, 1, 3, \dots\}$$

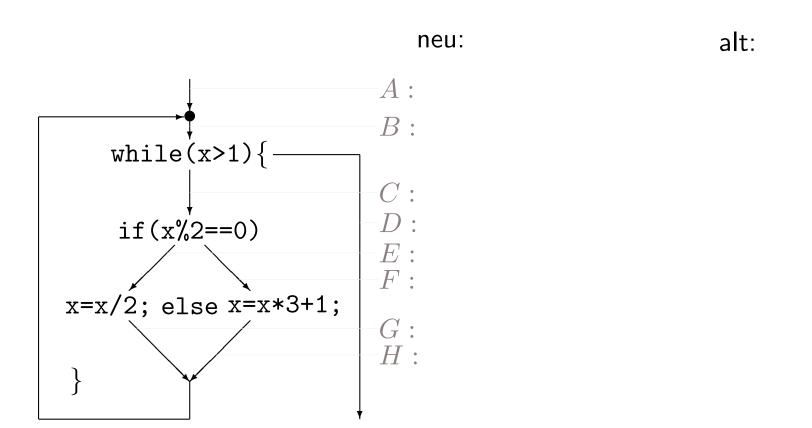


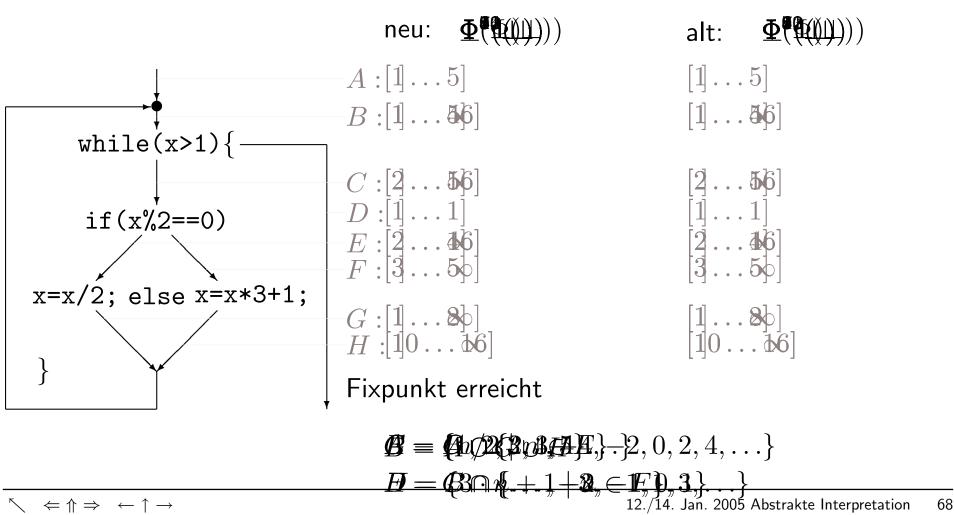


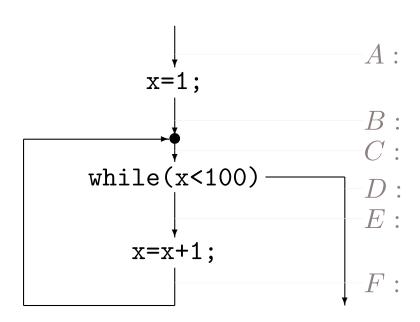
68

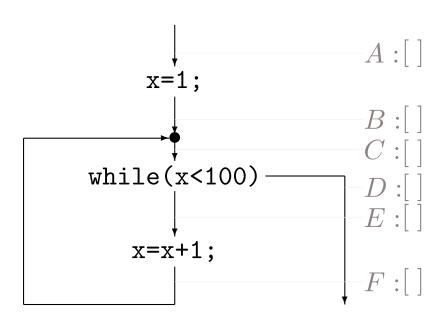




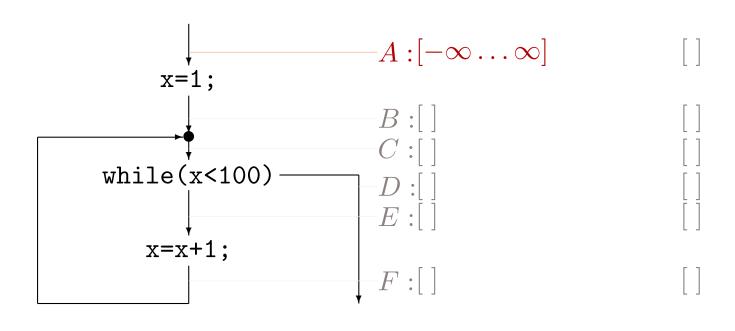


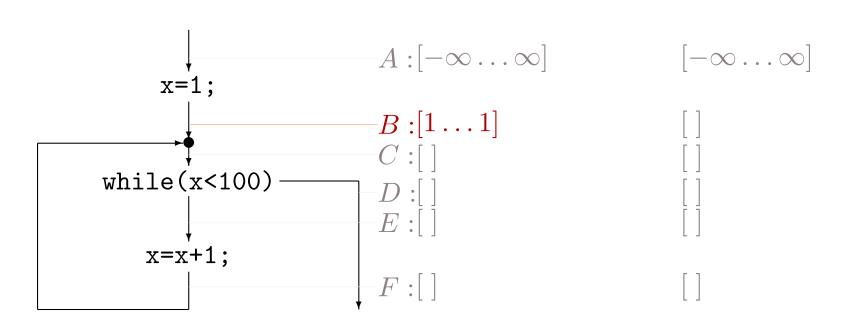




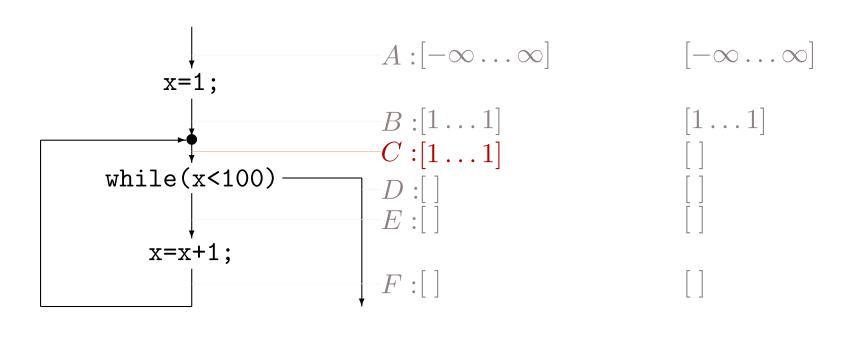


 $A = [-\infty \dots \infty]$ 

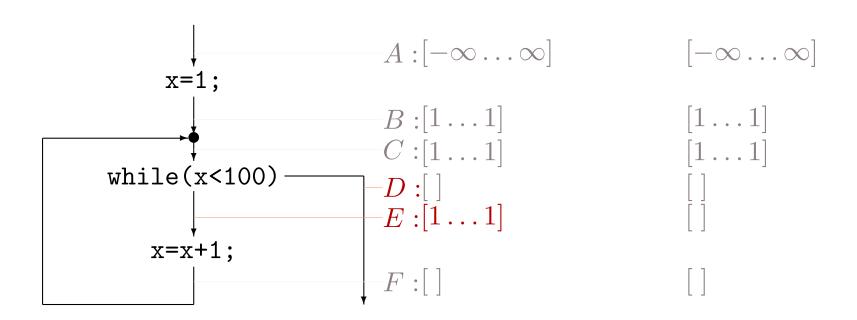




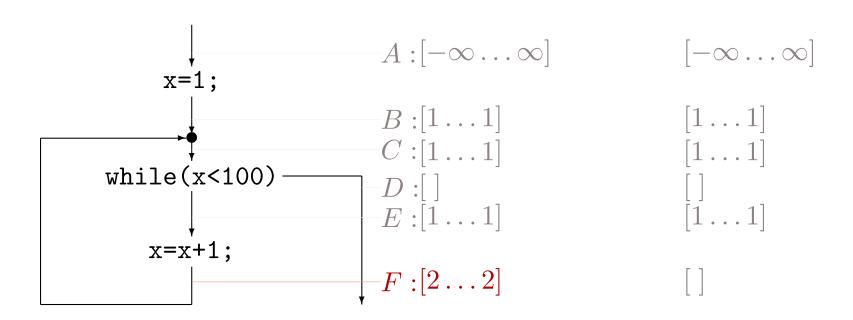
$$B = [1 \dots 1]$$



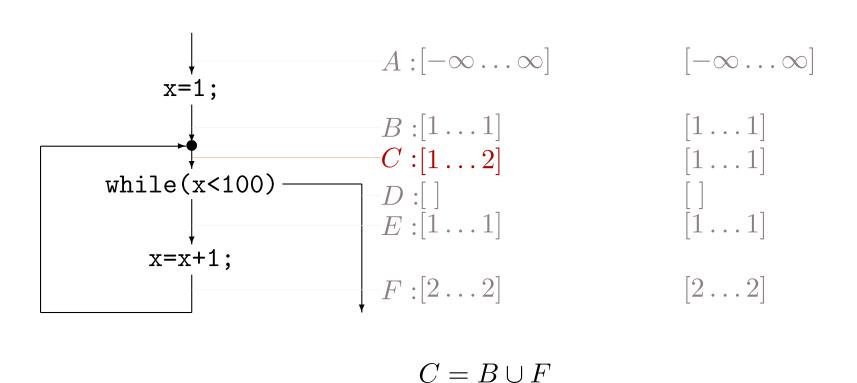
 $C = B \cup F$ 



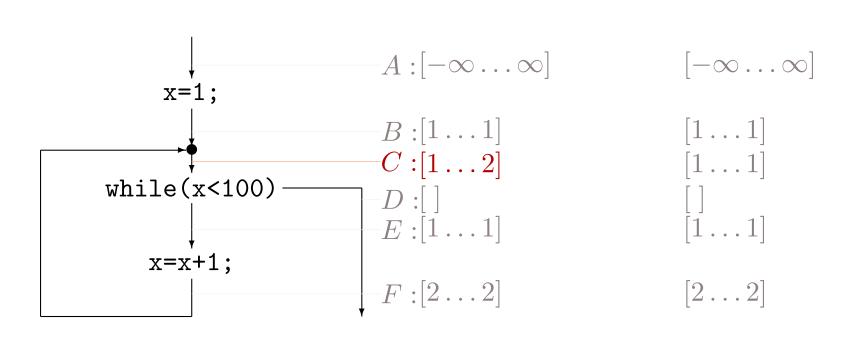
$$D=C\cap [100\ldots\infty]$$
  $E=C\cap [-\infty\ldots 99]$  12./14. Jan. 2005 Abstrakte Interpretation



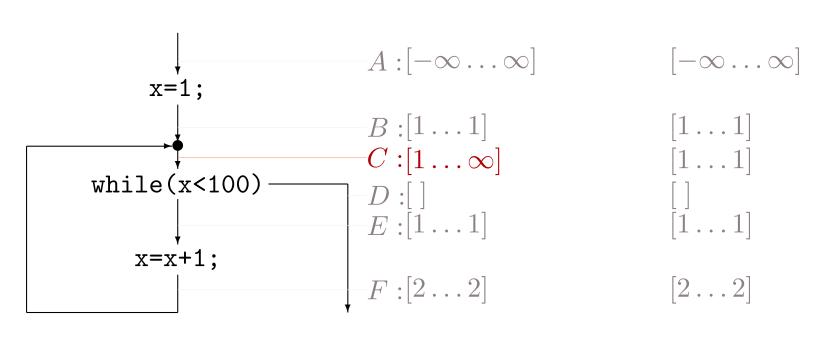
$$F = \{n+1 \mid n \in E\}$$



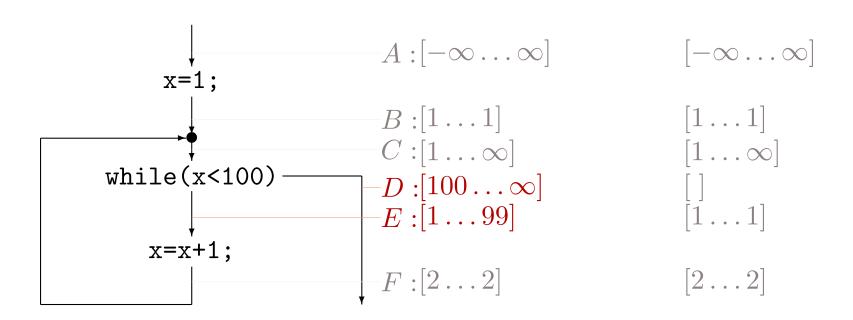
 $[1 \dots 1] \nabla [1 \dots 2] = [1 \dots \infty]$ 



$$^{\nwarrow}$$
  $\Leftarrow \uparrow \Rightarrow \leftarrow \uparrow \rightarrow$ 

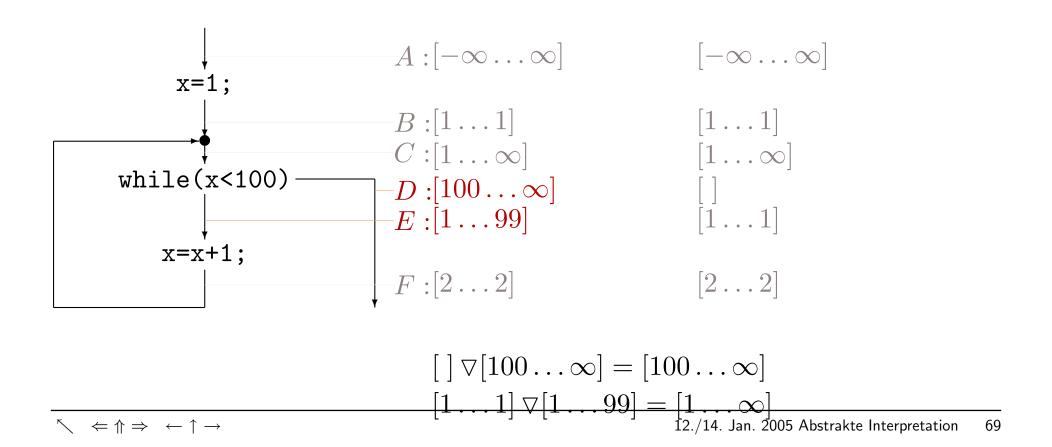


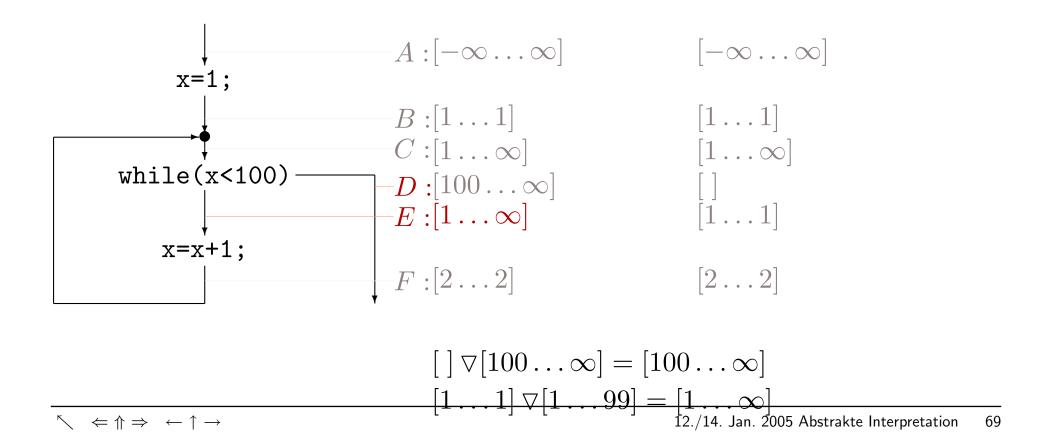
$$[1\dots 1] \, \nabla [1\dots 2] = [1\dots \infty]$$



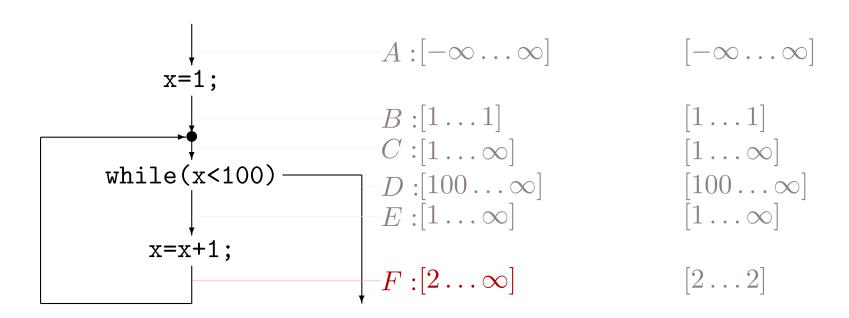
$$D = C \cap [100 \dots \infty]$$

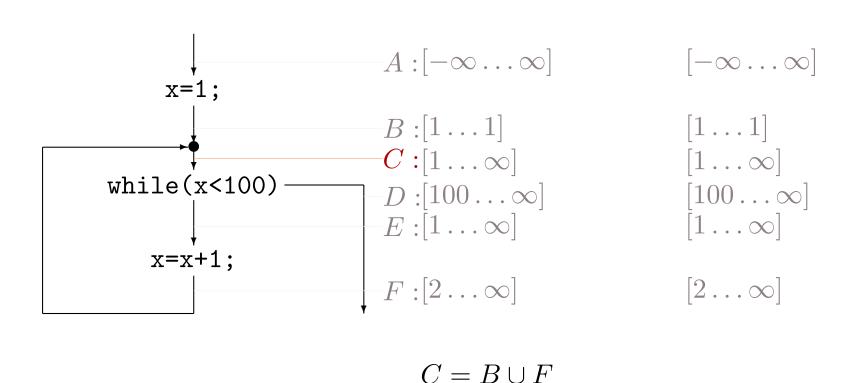
$$E = C \cap [\infty \quad 00]$$

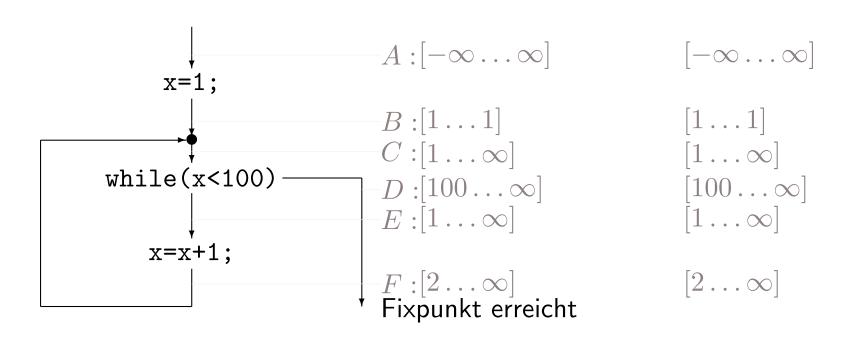


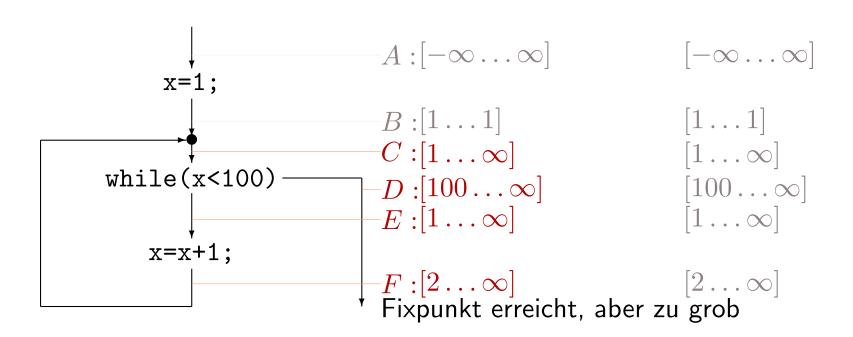


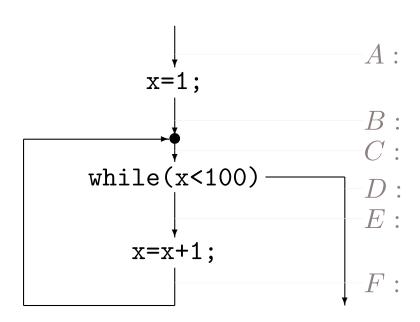
 $F = \{n+1 \mid n \in E\}$ 

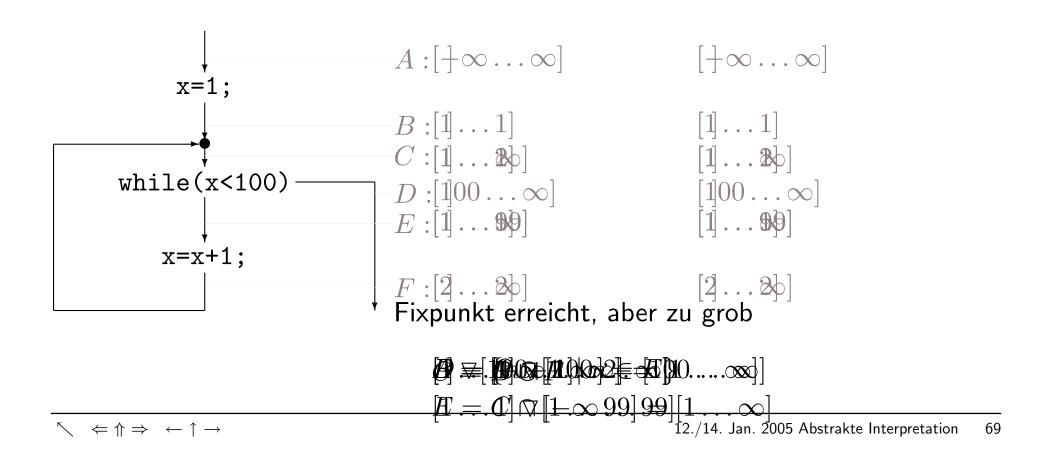












# Genauigkeitssteigerung durch nachgeschaltetes Narrowing

Sei  $\triangle: M' \times M' \longrightarrow M'$  eine Operation, so daß  $y' \sqsubseteq' x' \triangle y' \sqsubseteq' x'$  für alle  $y' \sqsubseteq' x'$  gilt, und so daß für jede absteigende Kette  $x_1' \sqsubseteq' x_2' \sqsubseteq' \dots \sqsubseteq' x_i' \sqsubseteq' \dots$  die Kette  $x_1' \sqsubseteq' x_1' \triangle x_2' \sqsubseteq' \dots \sqsubseteq' x_1' \triangle \dots \triangle x_i' \sqsubseteq' \dots$  irgendwann stationär wird.

Definiere  $\Phi'''(x') = x' \triangle \Phi'(x')$ .

Dann ist die Kette  $X'' \supseteq' \Phi'''(X'') \supseteq' \Phi'''(\Phi'''(X'')) \supseteq' \ldots \supseteq' \Phi'''^i(X'') \supseteq' \ldots$  irgendwann stationär und ihre größte untere Schranke  $X''' = \prod' \{X'', \Phi'''(X''), \Phi'''(\Phi'''(X'')), \ldots\}$  ist immer noch eine obere Abschätzung für den kleinsten Fixpunkt X' von  $\Phi'$ .

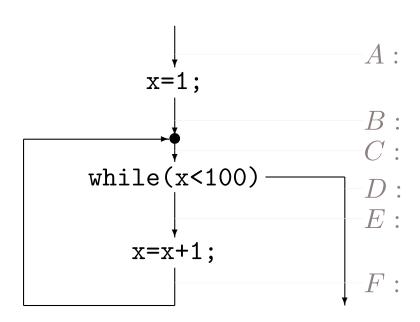
Definiere eine norrowing-Operation auf Intervallen durch

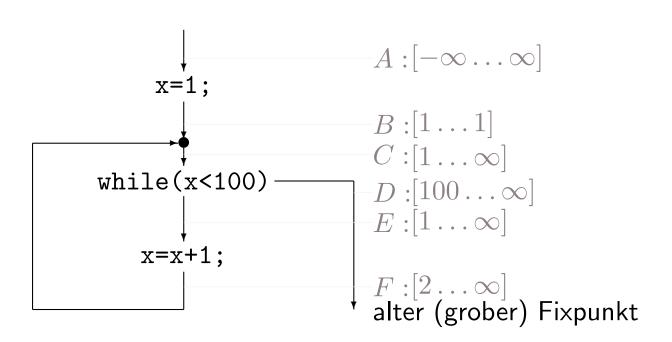
$$[l\dots u] \, \triangle[\,\,] = [\,\,] \, \triangle[l\dots u] = [l\dots u] \text{ sowie } [l_1\dots u_1] \, \triangle[l_2\dots u_2] = [l\dots u] \text{ mit } l_1 = -\infty$$
 
$$l = \left\{ \begin{array}{ll} l_2 & \text{für} & l_1 = -\infty \\ l_1 & \text{sonst} \end{array} \right. \text{ und } u = \left\{ \begin{array}{ll} u_2 & \text{für} & u_1 = +\infty \\ u_1 & \text{sonst} \end{array} \right.$$

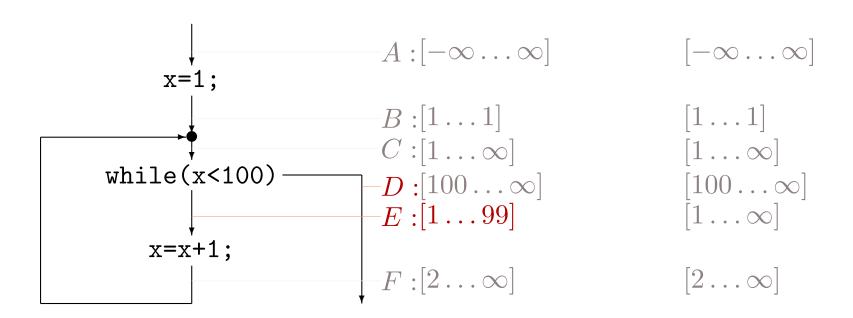
Anschaulich: Schlechteste Intervallgrenzen werden durch neue Werte ersetzt.

Z.B.

$$[1 \dots \infty] \triangle [1 \dots 16] = [1 \dots 16],$$
  
 $[1 \dots 16] \triangle [1 \dots \infty] = [1 \dots 16].$   
 $[1 \dots 16] \triangle [1 \dots 17] = [1 \dots 16].$ 



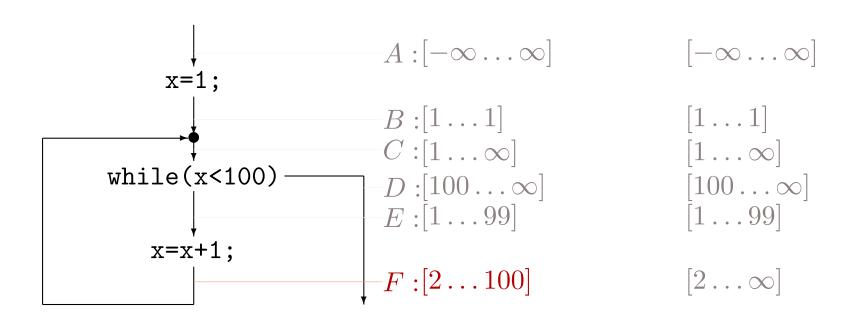




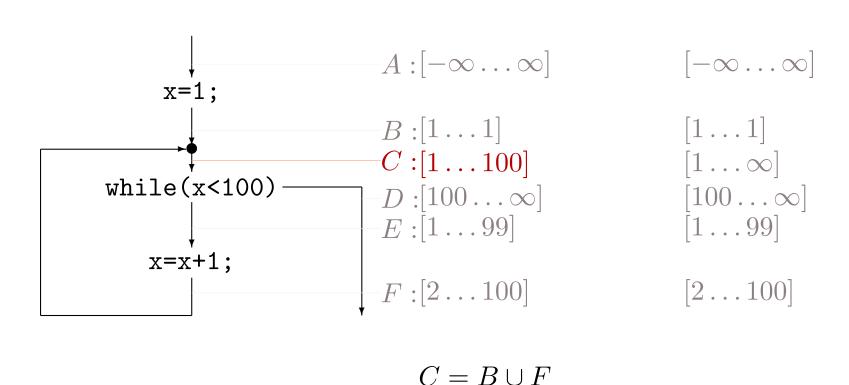
$$D = C \cap [100 \dots \infty]$$

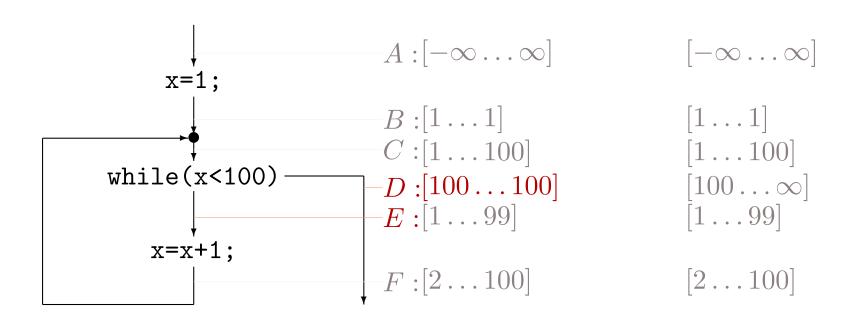
$$E = C \cap [-\infty \dots 99]$$

 $F = \{n+1 \mid n \in E\}$ 



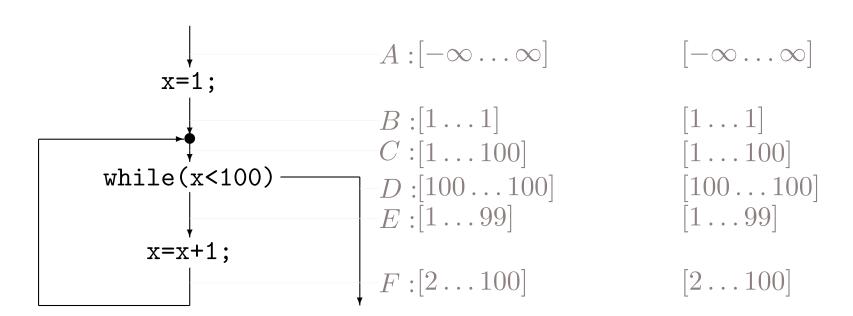
$$^{\nwarrow}$$
  $\Leftarrow \uparrow \Rightarrow \leftarrow \uparrow \rightarrow$ 

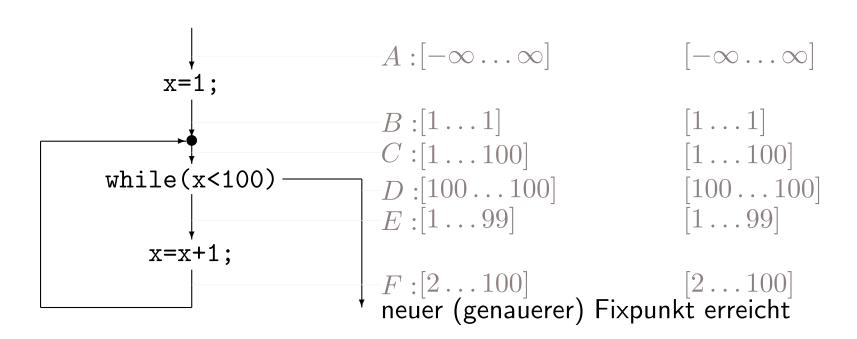


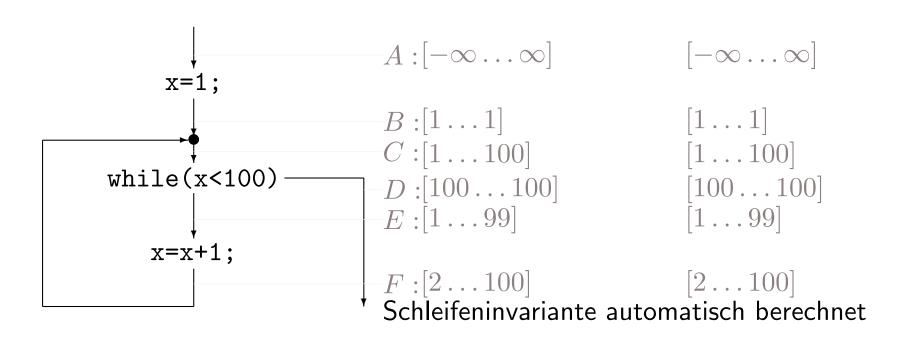


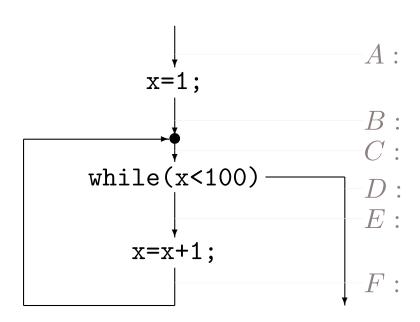
$$D = C \cap [100 \dots \infty]$$

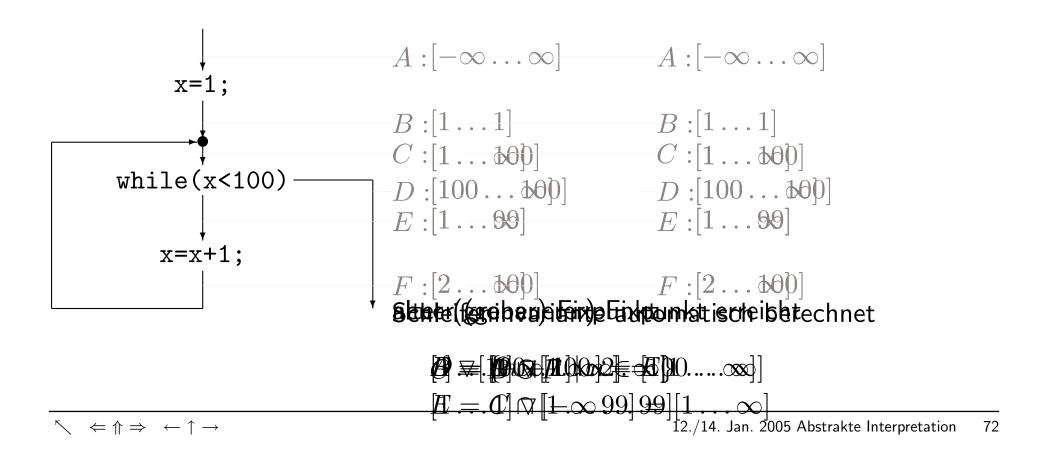
$$E = C \cap [-\infty \dots 99]$$











# Widening und Narrowing: Übersicht

# Anforderungen: Aussagekraft

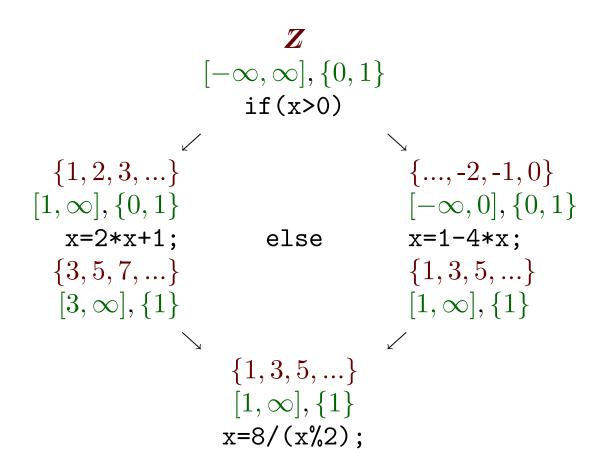
Die abstrakten Mengen sollen genügend präzise sein, um die (benutzergegebenen) Eigenschaften an den Programmpunkten entscheiden zu können.

Ein Indiz für zusätzliche Ungenauigkeit ist das Auftreten von Abschätzungen bei der Berechnung von  $\Phi'$  aus  $\Phi$ .

Manchmal kann dann der abstrakte Verband so erweitert werden (\*), daß genügend Information für eine genaueres  $\Phi'$  zur Verfügung steht.

(\*) durch Bildung kartesischer Produkte aus bekannten Einzelverbänden

#### **Beispiel**



- Abstrakter Verband: nur Intervalle
- Information über Rest mod.2 geht verloren
- Nulldivision scheint möglich
- Abstrakter Verband: zusätzlich Reste mod. 2 (kartesisches Produkt)
- $\Phi'$  für then- und else- Anweisung präziser
- Nulldivision ausgeschlossen

# Aussagekraft

Für die praktische Anwendung der abstrakten Interpretation ist es hinderlich, wenn für jedes zu analysierende Programm zuerst ein eigener abstrakter Verband ausgewählt und ein eigenes  $\Phi'$  berechnet werden muß.

Das Werkzeug PolySpace, das von Cousot-Schülern entwickelt wurde, arbeitet vorwiegend mit dem Verband der n-dimensionalen Polyeder (s.u.).

Bei int-Variablen wird, soweit möglich, zusätzlich Restklasseninformation mitgeführt (rationale Kongruenzen, s.u.).

# **Polyeder-Verband**

Der Polyeder-Verband ergibt sich als Verallgemeinerung des Intervall-Verbands auf mehrere Variablen.

Ein Element im Polyeder-Verband wird dargestellt durch eine Konjunktion linearer Ungleichungen.

Dadurch lassen sich lineare Abhängigkeiten zwischen Variablen noch berücksichtigen.

(In den Beispielen treten nur zwei Variablen auf, um eine geometrische Veranschaulichung zu ermöglichen.

Im allgemeinen Fall wird mit n-dimensionalen Polyedern gerechnet.)

$$x=80-2*y;$$

$$y=80/(x+y);$$

$$x \in [0, 99] \land y \in [0, 40]$$
 $x=80-2*y;$ 
 $y=80/(x+y);$ 
 $0$ 
 $0$ 
 $0$ 
 $0$ 

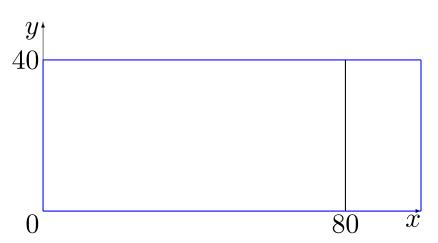
Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

$$x \in [0, 99] \land y \in [0, 40]$$
 $x=80-2*y;$ 
 $x \in [0, 80] \land y \in [0, 40]$ 
 $y=80/(x+y);$ 
 $0$ 
 $y=80/(x+y);$ 

Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

Das Ergebnis der Zuweisung x=80-2\*y; muß durch ein  $80 \times 40$ -Rechteck approximiert werden.

$$x \in [0,99] \land y \in [0,40]$$
  
 $\{0 \le x \le 99 \land 0 \le y \le 40\}$   
 $x=80-2*y;$   
 $x \in [0,80] \land y \in [0,40]$   
 $y=80/(x+y);$ 

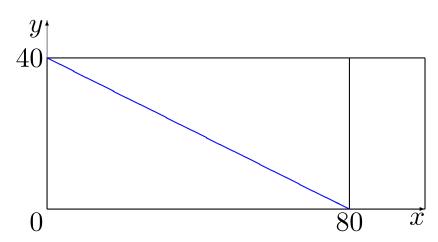


Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

Das Ergebnis der Zuweisung x=80-2\*y; muß durch ein  $80 \times 40$ -Rechteck approximiert werden.

Im Polyeder-Verband können beliebige konvexe n-Ecke dargestellt werden.

$$x \in [0,99] \land y \in [0,40]$$
  $\{0 \le x \le 99 \land 0 \le y \le 40\}$   $x=80-2*y$ ;  $x \in [0,80] \land y \in [0,40]$   $\{x=80-2y \land 0 \le y \le 40\}$   $y=80/(x+y)$ ;

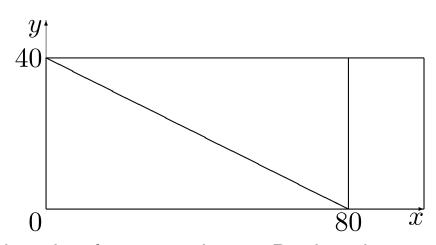


Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

Das Ergebnis der Zuweisung x=80-2\*y; muß durch ein  $80 \times 40$ -Rechteck approximiert werden.

Im Polyeder-Verband können beliebige konvexe n-Ecke dargestellt werden. Das Zuweisungsergebnis kann durch eine Linie exakt wiedergegeben werden.

$$x \in [0,99] \land y \in [0,40]$$
  $\{0 \le x \le 99 \land 0 \le y \le 40\}$  x=80-2\*y;  $x \in [0,80] \land y \in [0,40]$   $\{x = 80 - 2y \land 0 \le y \le 40\}$  y=80/(x+y);

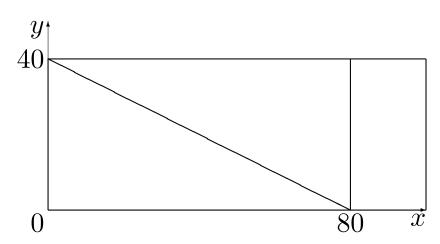


Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

Das Ergebnis der Zuweisung x=80-2\*y; muß durch ein  $80 \times 40$ -Rechteck approximiert werden.

Im Polyeder-Verband können beliebige konvexe n-Ecke dargestellt werden. Das Zuweisungsergebnis kann durch eine Linie exakt wiedergegeben werden. Eine Nulldivision in der folgenden Zuweisung kann damit ausgeschlossen werden.

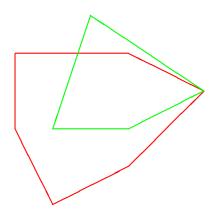
$$x \in [0,99] \land y \in [0,40]$$
  $\{0 \le x \le 99 \land 0 \le y \le 40\}$  x=80-2\*y;  $x \in [0,80] \land y \in [0,40]$   $\{x = 80 - 2y \land 0 \le y \le 40\}$  y=80/(x+y);

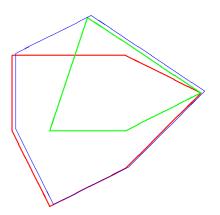


Im kartesischen Produkt aus zwei Intervallverbänden lassen sich nur Rechtecke darstellen.

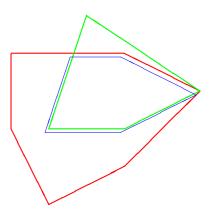
Das Ergebnis der Zuweisung x=80-2\*y; muß durch ein  $80 \times 40$ -Rechteck approximiert werden.

Im Polyeder-Verband können beliebige konvexe n-Ecke dargestellt werden. Das Zuweisungsergebnis kann durch eine Linie exakt wiedergegeben werden. Eine Nulldivision in der folgenden Zuweisung kann damit ausgeschlossen werden.



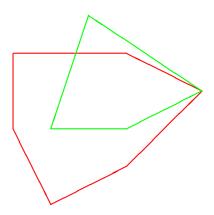


Die größte obere Schranke zweier Polyeder ist die konvexe Hülle ihrer Vereinigung.



Die größte obere Schranke zweier Polyeder ist die konvexe Hülle ihrer Vereinigung.

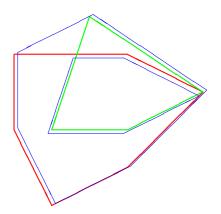
Die kleinste untere Schranke ist ihr Durchschnitt; er ist stets bereits konvex.



Die größte obere Schranke zweier Polyeder ist die konvexe Hülle ihrer Vereinigung.

Die kleinste untere Schranke ist ihr Durchschnitt; er ist stets bereits konvex.

Durchschnitt und konvexe Hülle der Vereinigung werden mit Hilfe geometrischer Algorithmen als neue Konjunktionen linearer Ungleichungen berechnet.



Die größte obere Schranke zweier Polyeder ist die konvexe Hülle ihrer Vereinigung.

Die kleinste untere Schranke ist ihr Durchschnitt; er ist stets bereits konvex.

Durchschnitt und konvexe Hülle der Vereinigung werden mit Hilfe geometrischer Algorithmen als neue Konjunktionen linearer Ungleichungen berechnet.

#### Polyeder-Verband: Analysebeispiel

```
i=2;
j=0;
while (...) {
       \{2j + 2 \le i \land 0 \le j\}
        if (...) {
               i=i+4;
                \{2j + 6 \le i \land 0 \le j\}
        } else {
               i=i+2;
               j=j+1;
                \{2j + 2 \le i \land 1 \le j\}
```

# Verband der Rationalen Kongruenzen

Verband der rationalen Kongruenzen:  $\{\bot, \top\} \cup (\mathbf{Q} \times \mathbf{Q})$ ,

wobei  $\langle p,q\rangle\in \mathbf{Q}\times\mathbf{Q}$  intuitiv für  $\{p+k\cdot q\mid k\in\mathbf{Z}\}=\gamma(\langle p,q\rangle)$  steht.

Z.B. 
$$\gamma(\langle 1.0, 10 \rangle) = \{\dots, -19, -9, 1, 11, 21, 31, \dots\}$$
 
$$\gamma(\langle 0.5, 1.0 \rangle) = \{\dots, -1.5, -0.5, 0.5, 1.5, 2.5, \dots\}$$
 
$$\gamma(\langle 0.0, 0.1 \rangle) = \{\dots, -0.2, -0.1, 0, 0.1, 0.2, 0.3, \dots\}$$

Dieser Verband ist unendlich groß und erfüllt nicht die ACC-Bedingung, so daß mit Widening und Narrowing gearbeitet werden muß.

Das Gleiche gilt für den Polyeder-Verband.

Cousot, Cousot 1992 gibt für jeden der beiden Verbände ein geeignetes Widening und Narrowing an.

# Statische Analyse mit PolySpace

(Werbefolien für Industriekunden)

- Code wird analysiert, nicht ausgeführt
- prüft minimale semantische Kriterien (z.B. NULL-Pointer-Zugriffe)
- vollautomatisch, Rechenzeit im Stundenbereich

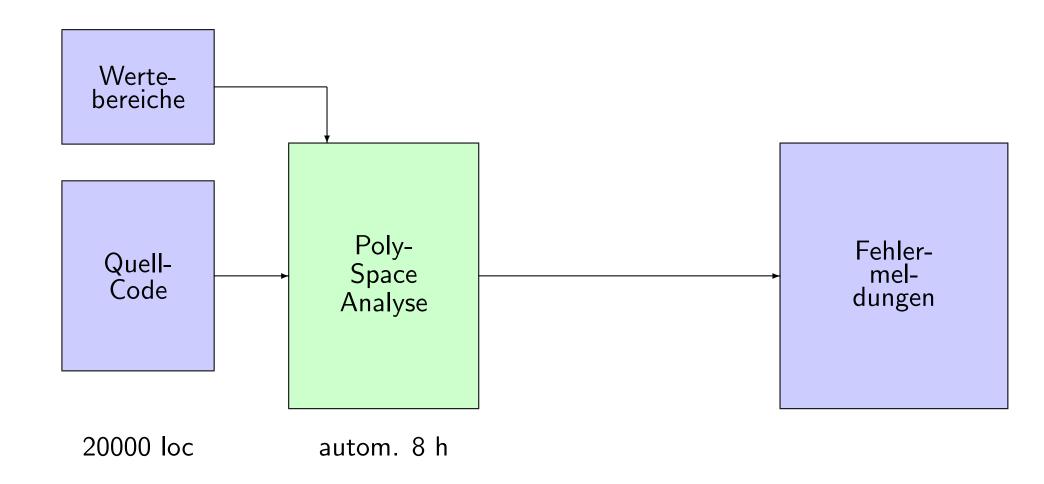
# **Statische Analyse**

- unabhängig von Programm-Funktionalität
- benötigt keine Anforderungsspezifikation
- Aufwand und Nutzen zwischen Compiler und Test
- führt zu robusterem Code

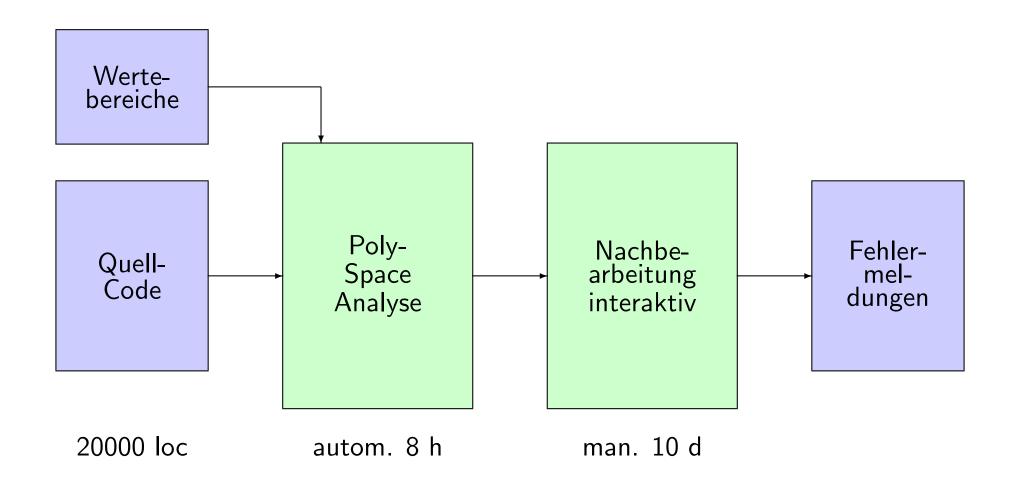
## **Statische Analyse**

- Grundlage: Wertebereichsanalyse
- an ca. 0.1% der Codestellen Fehler gefunden
- an ca. 10% der Codestellen Fehler nur vermutet
- manuelle Nachkontrolle unabhängig von Entwicklern

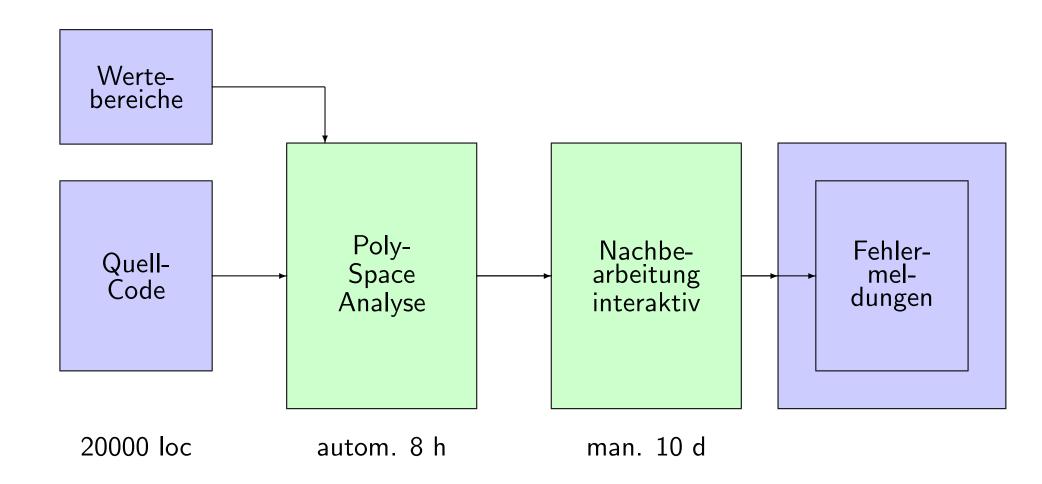
# Vorgehen



## Vorgehen



## Vorgehen



#### **Untersuchte Fehlerarten**

- Uninitialisierte Variablen
- Zugriff über ungültigen Zeiger
- Unerreichbarer Code
- Wertebereichsüberlauf

- Indexüberlauf
- Nulldivision
- Endlosschleife
- Endlosrekursion

```
int a[10];
...
for (i=0; i<10; ++i)
    a[i] = a[i+1];</pre>
```

```
int a[10];
...
for (i=0; i<10; ++i)
    a[i] = a[i+1];</pre>
```

```
int a[10];
...
for (i=0; i<10; ++i)
    a[i] = a[i+1];</pre>
```

```
int a[10];
...
for (i=0; a[i]>0; ++i)
    a[i] = a[i+1];
```

Manuelle Nachkontrolle notwendig

- Manuelle Nachkontrolle notwendig
- Programmablauf muß lokal verstanden werden

```
int a[10];
...
for (i=0; a[i]>0; ++i)
    a[i] = a[i+1];
```

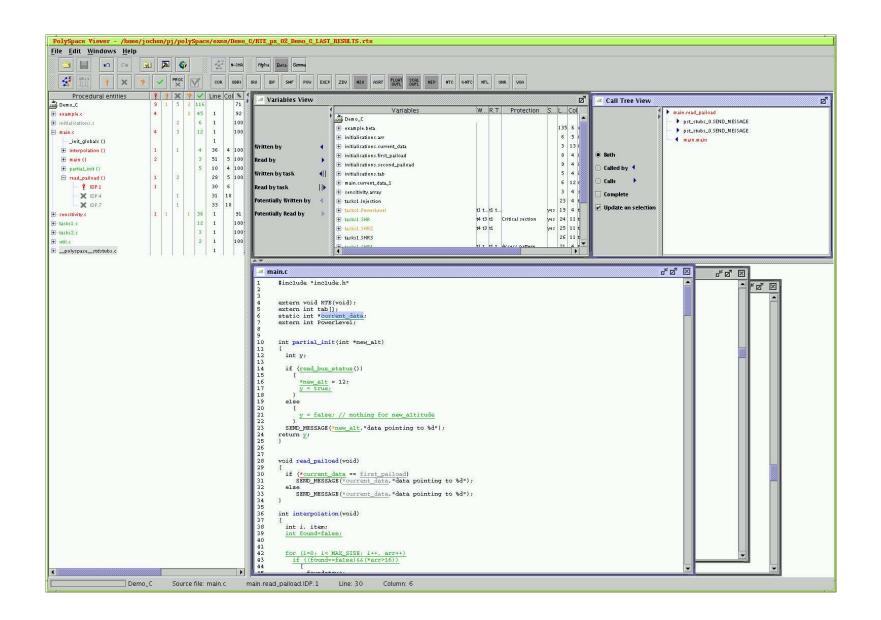
- Manuelle Nachkontrolle notwendig
- Programmablauf muß lokal verstanden werden

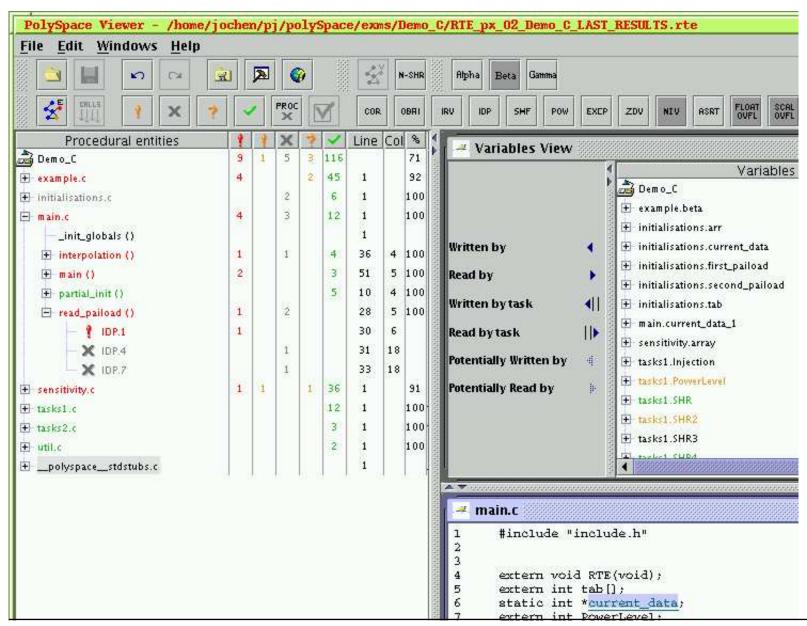
#### Uninitialisierte Variablen

```
int search(int a[],int f,int t,int v) {
    while (f < t) {
        m = (f + t) / 2;
        if (v<a[m]) t=m; else if ...
    }
    return m;
}</pre>
```

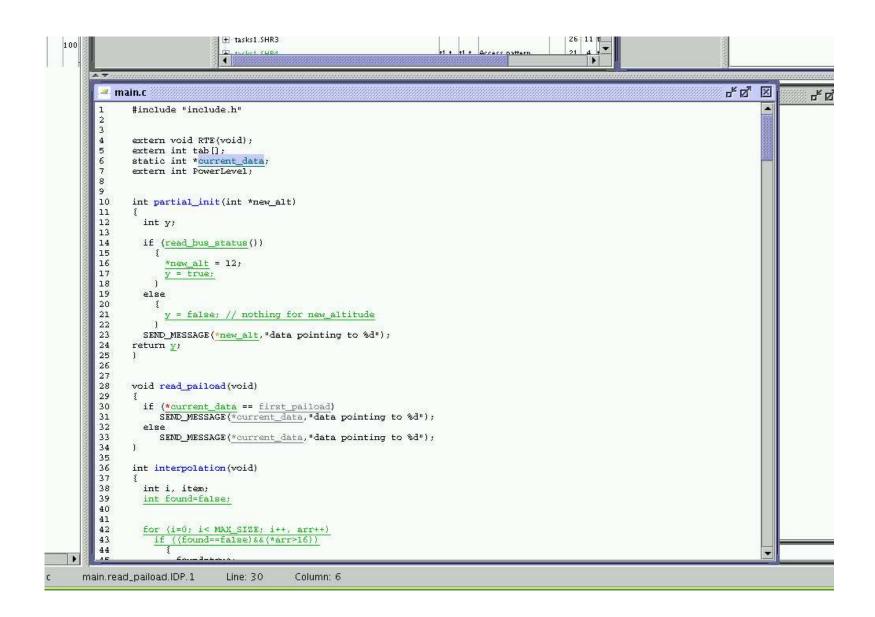
#### Uninitialisierte Variablen

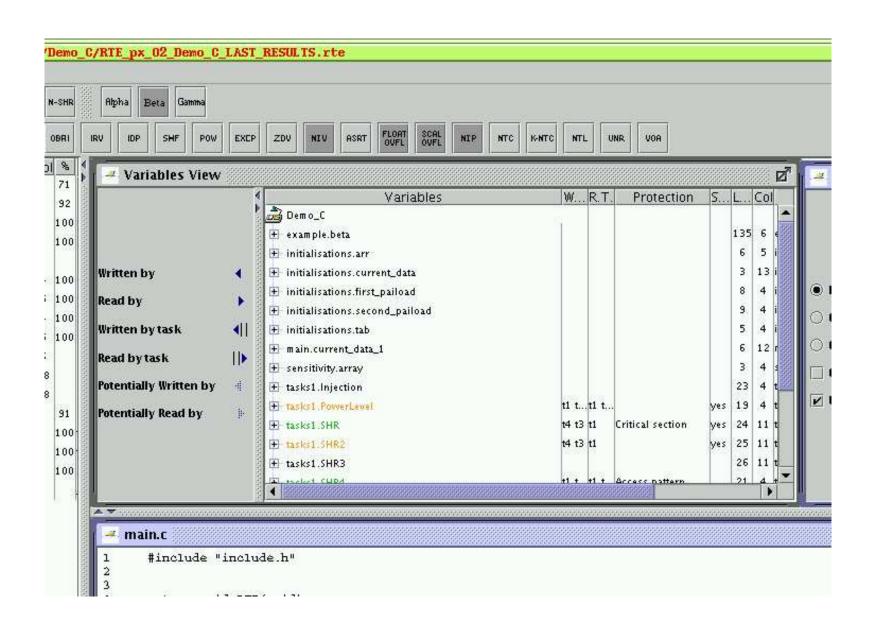
# Werkzeugoberfläche

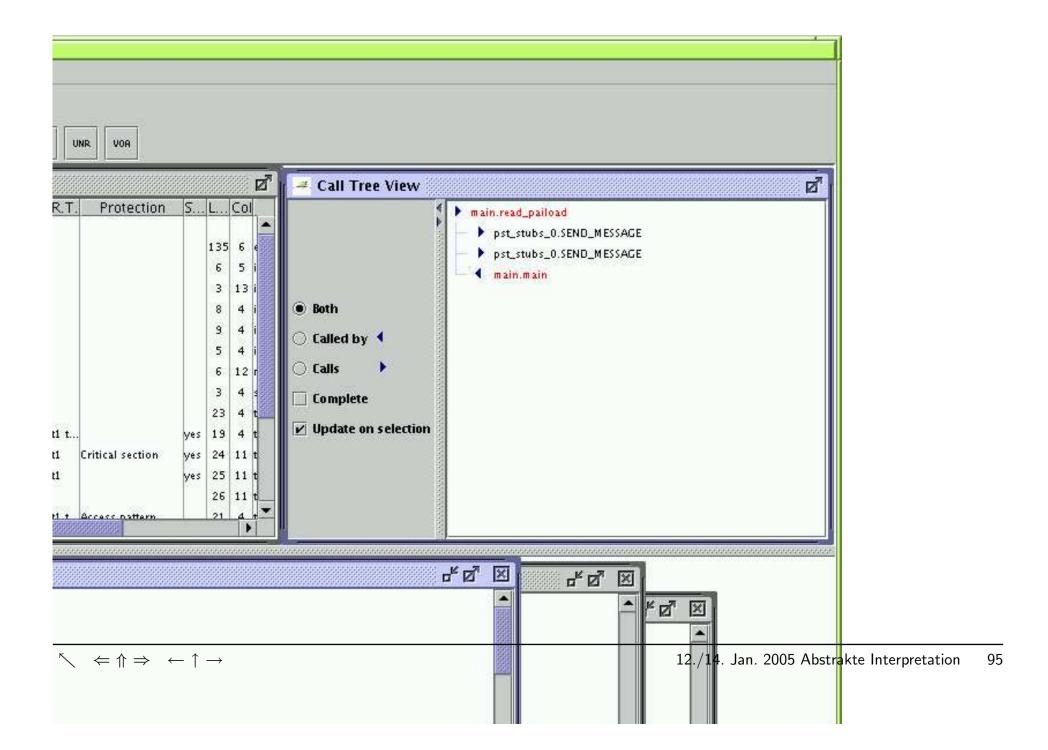


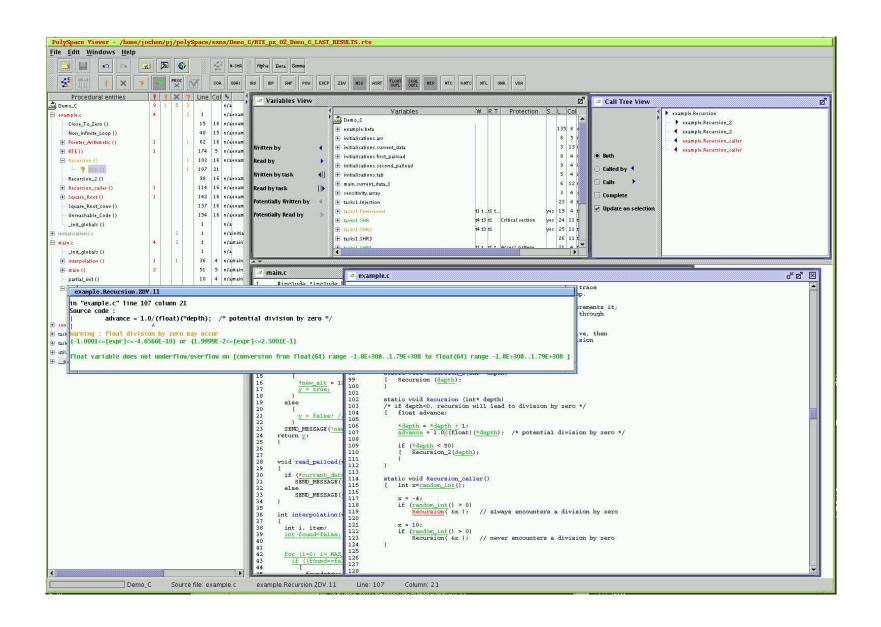


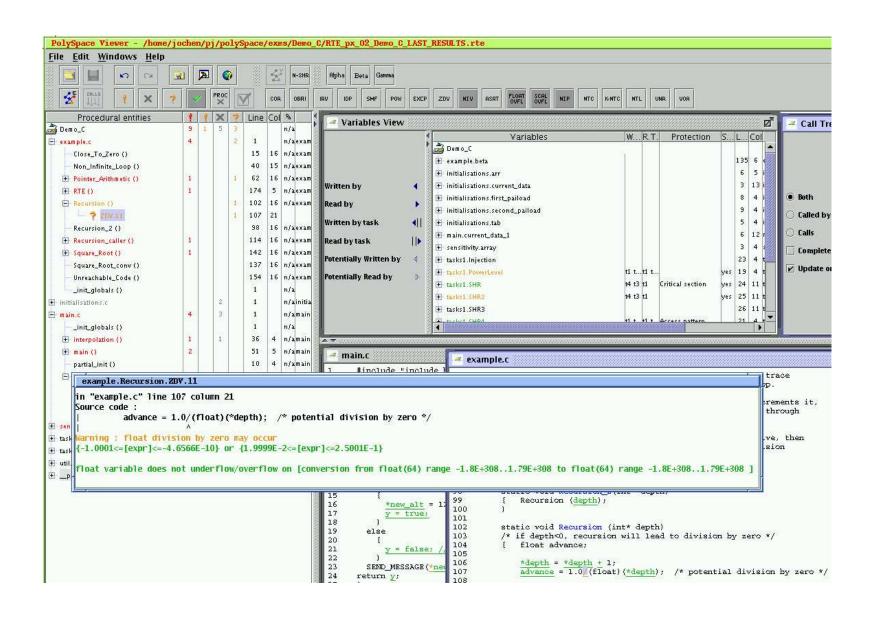
92











## Ergänzungen

Abstrakte Interpretation für nicht imperative Programme:

Voraussetzung ist, daß sich aus einem gegebenen Programm der Operator  $\Phi$  bestimmen läßt.

(Auch bei imperativen Programmen sind noch viele Verallgemeinerungen möglich. Wir haben nur ein spezielles  $\Phi$  betrachtet, nämlich das, das zur collecting semantics führt. Z.B. Datenflußanalyse ist mit einem anderen  $\Phi$  möglich.)

# Anwendungen in Übersetzerbau und Verifikation

- Unbenutzte Programmabschnitte
- Feldgrenzenüberwachung
- Vorzeichen
- Datenflußanalyse
- Auslagerung aus Schleifen

# Anwendungen in Übersetzerbau und Verifikation

- Reference counting
- Slicing
- Erzeugung von Schleifeninvarianten
- Terminierung von Schleifen
- Worst case execution time

## Werdegang

- 1960er Jahre: "Symbolic execution" (P. Naur)
- danach: viele ad-hoc-Verfahren, mehrere fehlerhaft
- 1976 / 1977: Cousot, Cousot Allgemeine Methodik für abstrakte Interpretation von Flowchart-Programmen
- danach: Erweiterung auf funktionale und logische Programme
- 1990er Jahre: Kombination mit Model-Checking

#### Literatur

#### References

- [CC76] P. Cousot and R. Cousot. Static determination of dynamic properties of programs. In *Proc. 2nd Int. Symp. on Programming*, pages 106–130, Paris, 1976. Dunot.
- [CC92] P. Cousot and R. Cousot. Comapring the Galois connection and widening / narrowing approaches to abstract interpretation. In Maurice Bruynooghe and Martin Wirsing, editors, Proc. 4th Int. Symp. on Programming Language Implementation and Logic Programming (PLILP), volume 631 of LNCS, pages 269–296, Heidelberg, Aug. 1992. Springer.