4. Abstrakte Interpretation

Ziel: Entwickle Programmanalysen,
die die Semantik korrekt approximieren
↳ Die bisherigen Datenflussanalysen haben
die Semantik von Programmen nicht berücksichtigt.
(bestenfalls intuitiv)

Idee: · Führe das Programm auf abstrakten Werten aus.
↳ Beispiel $\mathcal{P}(\{-, 0, +\})$ statt $\mathbb{Z}$.
↳ Berücksichtige alle konkreten Eingaben.
· Ersetze konkrete Operationen durch abstrakte Operationen
↳ Es müssen alle konkreten Werte berücksichtigt werden,
die durch abstrakten Wert dargestellt sind.
↳ Beispiel $op(x) = x - 2$.
Dann $op^{\#}(\{+\}) = \{-, 0, +\}$
$op^{\#}(\{0\}) = \{-\} = op^{\#}(\{-\})$.

Vorteile: · Mächtigkeit
↳ Abstrakte Interpretation ist unabhängig von Kontrollflussgraph
↳ Funktioniert für viele Klassen von Programmen
(if/while, parallel, Objekt-orientiert, Aktoren, funktional, ...)
· Korrektheit
↳ Durch die Theorie garantiert.
· Balance zwischen Präzision und Komplexität der Analyse
↳ Variierbar durch Granularität der abstrakten Domäne.

Nachteil: Komplexität
↳ Bei abstrakter Interpretation oft höher
als bei Datenflussanalysen.

4.1 Galois-Verbindungen $(L \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} M)$

Ziel: · Beschreibe geeignete Abstraktionsfunktionen $\alpha : L \to M$
die jeden konkreten Wert auf einen abstrakten Wert abbilden.

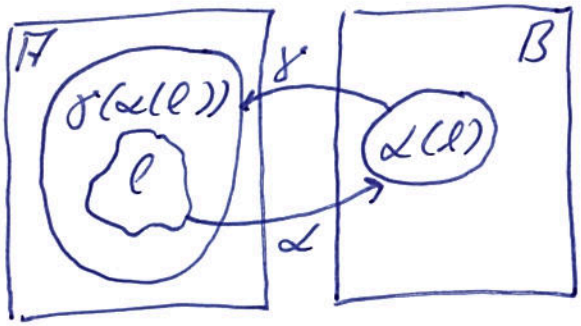
- Definiere neben α auch Konkretisierungsfunktion $\gamma: M \to L$, die jedem abstakten Wut alle konkreten Werte zuordnet, für die α steht.

Definition (Galois-Verbindung):
Seien $(L, \leq_L)$ und $(M, \leq_M)$ vollständige Verbände.
Ein Paar (α, γ) von monotonen Funktionen $\alpha: L \to M$ und $\gamma: M \to L$ heißt Galois-Verbindung, falls

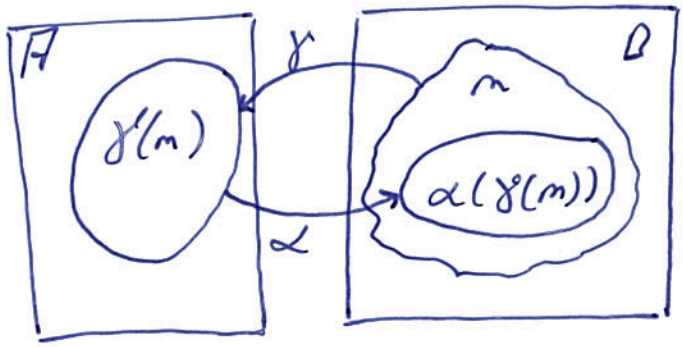
(G1) $\forall l \in L: \; l \leq_L \gamma(\alpha(l))$

(G2) $\forall m \in M: \; \alpha(\gamma(m)) \leq_M m$.

Anschaulich:
Seien $L = \mathcal{P}(\Pi)$ mit Π = Menge konkreter Werte
und $M = \mathcal{P}(B)$ mit B = Menge abstakte Werte.



(G1) $l \subseteq \gamma(\alpha(l))$,
α erzeugt Überapproximation



(G2) $\alpha(\gamma(m)) \subseteq m$,
kein Präzisionsverlust durch Abstahtion nach Konkretisierung

Typischerweise:
$l \neq \gamma(\alpha(l))$ aber $m = \alpha(\gamma(m))$.

Satz (Eigenschaften von Galois-Verbindungen I):
Sei (α, γ) eine Galois-Verbindung mit $\alpha: L \to M$ und $\gamma: M \to L$.
Seien ferner $l \in L$ und $m \in M$.
(1.) $\alpha(l) \leq_M m$ gdw. $l \leq_L \gamma(m)$
(Aus dieser Äquivalenz folgt auch wieder, dass es sich um eine Galois-Verbindung handelt).

(2.a) Die Konkretisierung γ ist eindeutig durch α bestimmt:

$$\gamma(m) = \bigsqcup \{ l \in L \mid \alpha(l) \leq_M m \}.$$

(2.b) Die Abstraktion α ist eindeutig durch γ bestimmt:

$$\alpha(l) = \bigsqcap \{ m \in M \mid l \leq_L \gamma(m) \}$$

Beweis:

(1.) Gelte $\alpha(l) \leq_M m$.
Dann folgt
$l \leq_L \gamma(\alpha(l)) \leq_L \gamma(m)$
(G1) (Monotonie γ)

Ähnlich sei $l \leq_L \gamma(m)$.
Dann
$\alpha(l) \leq_M \alpha(\gamma(m)) \leq_M m$.
(Monotonie α) (G2)

(2.a) Zeige $\gamma(m) = \bigsqcup \{ l \in L \mid \alpha(l) \leq_M m \}$ für jede Galois-Verbindung (α, γ).
Zeige dazu $\leq_L$ und $_L\geq$.

Zu $_L\geq$:
Falls $\alpha(l) \leq_M m$, dann folgt $l \leq_L \gamma(m)$ mit (1.).
Damit ist
$$\gamma(m) \;{}_L\!\geq \bigsqcup \{ l \in L \mid \alpha(l) \leq_M m \}.$$

Zu $\leq_L$:
Da $\alpha(\gamma(m)) \leq_M m$, gilt
$$\gamma(m) \in \{ l \in L \mid \alpha(l) \leq_M m \}.$$
Also
$$\gamma(m) \leq_L \bigsqcup \{ l \in L \mid \alpha(l) \leq_M m \}.$$
□

Satz (Eigenschaften von Galois-Verbindungen II)
Sei (α, γ) eine Galois-Verbindung mit $\alpha : L \to M$ und $\gamma : M \to L$.
Seien sowie $L' \subseteq L$ und $M' \subseteq M$.
(3.a) α ist vollständig distributiv (auch vollständig additiv genannt):
$$\alpha(\bigsqcup L') = \bigsqcup \alpha(L').$$
(3.b) γ ist vollständig multiplikativ:
$$\gamma(\bigsqcap M') = \bigsqcap \gamma(M').$$

(4.a) Zu jeder vollständig additiven Funktion $\alpha: L \to M$ gibt es eine Funktion $\gamma: M \to L$, so dass (α, γ) eine Galois-Verbindung ist.

(4.b) Ebenso gibt es zu jeder vollständig multiplikativen Funktion $\gamma: M \to L$ eine Funktion $\alpha: L \to M$, so dass (α, γ) eine Galois-Verbindung ist.

Beweis:

(3.a) Um $\alpha(\bigsqcup L') = \bigsqcup \alpha(L')$ zu zeigen, zeige $\leq_M$ und $\geq_M$.

Zu $\geq_M$:

Für $l \in L'$ gilt $l \leq_L \bigsqcup L'$.
M.t der Monotonie von α folgt

$$\alpha(l) \leq_M \alpha(\bigsqcup L').$$

Da die Ungleichung für alle $l \in L'$ gilt, folgt

$$\bigsqcup \alpha(L') = \bigsqcup \{\alpha(l) \mid l \in L'\} \leq_M \alpha(\bigsqcup L').$$

Zu $\leq_M$:

Um $\alpha(\bigsqcup L') \leq_M \bigsqcup \alpha(L')$ zu zeigen, zeige

$$\bigsqcup L' \leq_L \gamma(\bigsqcup \alpha(L')).$$

Dann folgt die gewünschte Ungleichung mit Eigenschaft (1.). □

4.2 Konstruktion von Galois-Verbindungen

Ziele: (1) Definiere zwei elementare Galois-Verbindungen
(2) Definiere Galois-Verbindungen mittels Existenzfunktionen
(3) Komponiere bestehende Galois-Verbindungen zu neuen.

1.a) Intervallabstraktion:

- Intervallabstraktionen erlauben Aussagen über die absolute Größe von Datenwerten.
- Aber es lässt sich mit abstrakten Intervallwerten nur unpräzise rechnen.

Sei $L = (\mathcal{P}(\mathbb{Z}), \subseteq)$ die konkrete Domäne der Teilmengen von $\mathbb{Z}$.

Sei $M = \left(\left((\mathbb{Z} \cup \{-\infty\}) \times (\mathbb{Z} \cup \{+\infty\}) \right) \cup \{\emptyset\}, \subseteq \right)$
die abstrakte Domäne der Intervalle.

Definiere $\alpha : L \to M$ mittels

$$\alpha(Z) := \begin{cases} \emptyset, & \text{falls } Z = \emptyset \\ [\sqcap Z, \sqcup Z], & \text{sonst.} \end{cases}$$

$\gamma : M \to L$ mittels

$$\gamma(I) := \begin{cases} \emptyset, & \text{falls } I = \emptyset \\ \{z \in \mathbb{Z} \mid z_1 \leq z \leq z_2\}, & \text{falls } I = [z_1, z_2]. \end{cases}$$

Zum Beispiel gilt

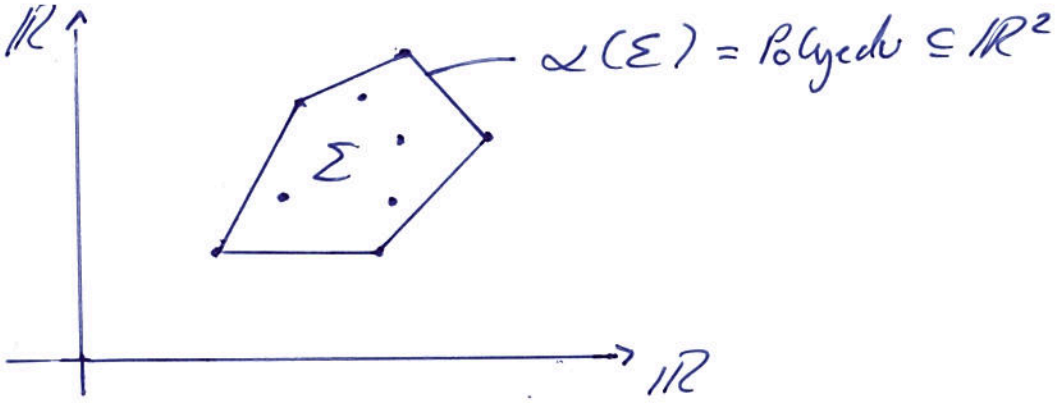
- $\gamma(\alpha(\{1, 3, 5, \ldots\})) = \gamma([1, +\infty]) = \{1, 2, 3, 4, \ldots\} \supseteq \{1, 3, 5, \ldots\}$
- $\alpha(\gamma([-1, 1])) = \alpha(\{-1, 0, 1\}) = [-1, 1].$

In der Praxis ist

$$M = \left(\left((\{k_1, \ldots, k_n\} \cup \{-\infty\}) \times (\{k_1, \ldots, k_n\} \cup \{+\infty\}) \right) \cup \{\emptyset\}, \subseteq \right)$$

Verallgemeinerung: Approximation durch konvexe Polyeder

Sei $\Sigma \subseteq \mathbb{R}^{|Vars|}$ eine Menge von Variablenbelegungen (mit reellen Zahlen).
M.i. $n = |Vars|$ lässt sich Σ als Punktmenge in $\mathbb{R}^n$ auffassen.
Man approximiert Σ durch das kleinste konvexe Polyeder, das alle Punkte enthält.



Polyeder lassen sich endlich darstellen,
zum Beispiel, indem man Eckpunkte aufzählt.