Beiträge zu praktikabler Prädikatenanalyse

Towards Practical Predicate Analysis

Philipp Wendler

8. Mai 2018

UNIVERSITÄT PASSAU
Fakultät für Informatik und Mathematik

SoSy-Lab
Software Systems

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

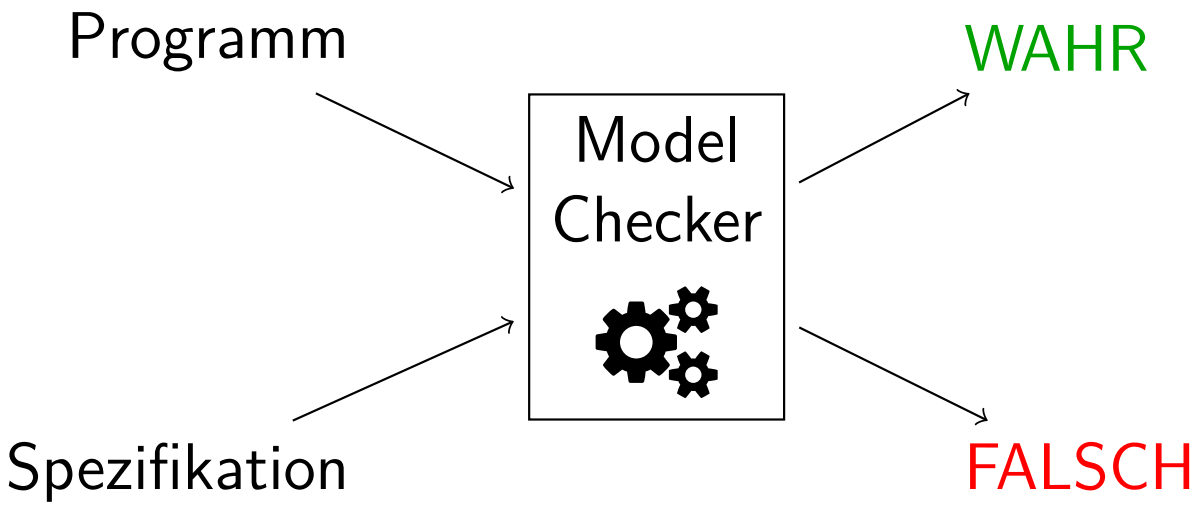
Wozu automatische Software-Verifikation?

- Software ist kritisch in der heutigen Welt
- Software hat Fehler
- Software ist zu komplex um alle Fehler manuell zu finden

Wozu automatische Software-Verifikation?

- Software ist kritisch in der heutigen Welt
- Software hat Fehler
- Software ist zu komplex um alle Fehler manuell zu finden

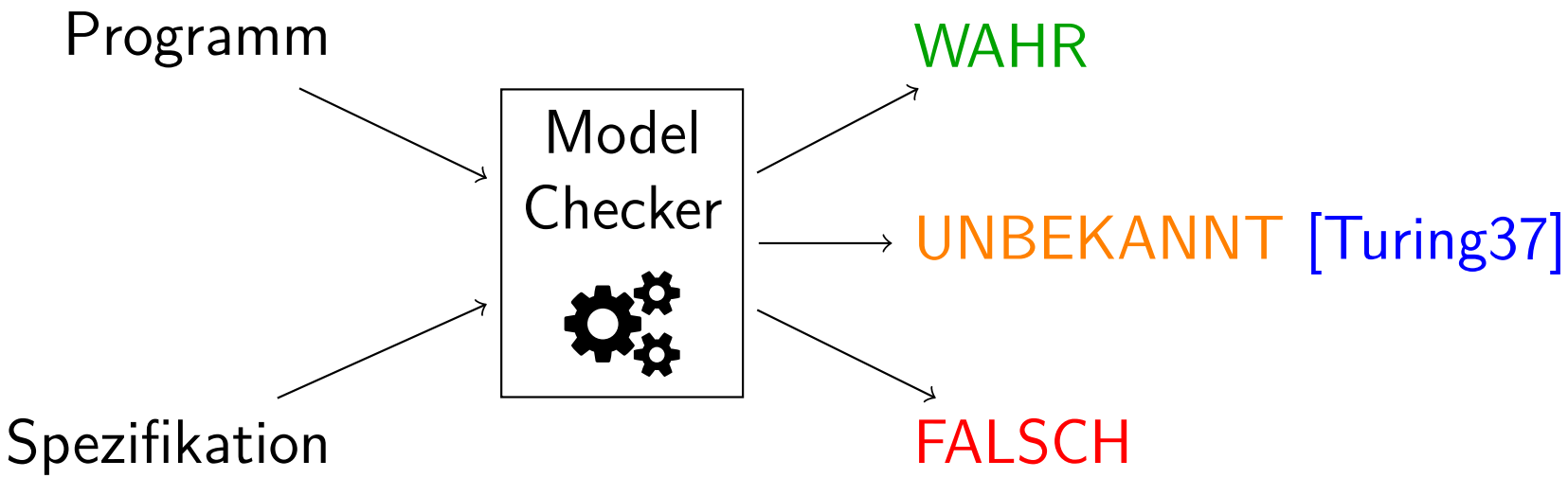
Lösung “Model-Checking”:



Wozu automatische Software-Verifikation?

- Software ist kritisch in der heutigen Welt
- Software hat Fehler
- Software ist zu komplex um alle Fehler manuell zu finden

Lösung “Model-Checking”:



Prädikatenanalyse

Viele Ansätze teilen sich diese Idee:

- Konvertiere Teile des Programms in Formeln der Prädikatenlogik erster Stufe (Erfüllbarkeit modulo Theorien, SMT)
  Beispiel: $x > 0 \wedge y < 10$
- Frage Solver nach Erfüllbarkeit

Wir nennen SMT-basierte Ansätze *Prädikatenanalyse*.

- Profitieren von den Fortschritten moderner Solver
- In der Praxis verwendet für Software-Verifikation

Existierende Prädikatenanalysen

- Prädikatenabstraktion
- IMPACT
- Bounded Model-Checking
- k-Induktion
- ...

Existierende Prädikatenanalysen und Tools

- Prädikatenabstraktion
  (Blast, Slam, ...)
- Impact
  (Impact, Wolverine, ...)
- Bounded Model-Checking
  (Cbmc, Esbmc, ...)
- k-Induktion
  (Esbmc, 2ls, ...)
- ...

Probleme mit dem Stand der Forschung (vorher)

Ansätze existieren isoliert

- Unterschiede und Gemeinsamkeiten schwer zu sehen
- Kernideen und Vorteile schwer zu verstehen
- Bremst Forschung aus

Probleme mit dem Stand der Forschung (vorher)

Ansätze existieren isoliert

- Unterschiede und Gemeinsamkeiten schwer zu sehen
- Kernideen und Vorteile schwer zu verstehen
- Bremst Forschung aus

Ansätze sind implementiert in einzelnen Tools

- . . . von unterschiedlicher Qualität (akademische Prototypen)
- Experimenteller Vergleich von Ansätzen schwierig
- Bremst Forschung und Einsatz in Praxis aus

Auf einer Tagung gestellte Frage

Ihre Vermutung:
Welcher Ansatz ist besser (löst mehr Verifikationsaufgaben)?

k-Induktion

Prädikatenabstraktion

Auf einer Tagung gestellte Frage

Ihre Vermutung:
Welcher Ansatz ist besser (löst mehr Verifikationsaufgaben)?

k-Induktion

Prädikatenabstraktion

Je nach Konfiguration der eine oder der Andere!

Technische Details (z. B. Wahl der SMT-Theorie)
beeinflussen Beurteilung der Algorithmen.

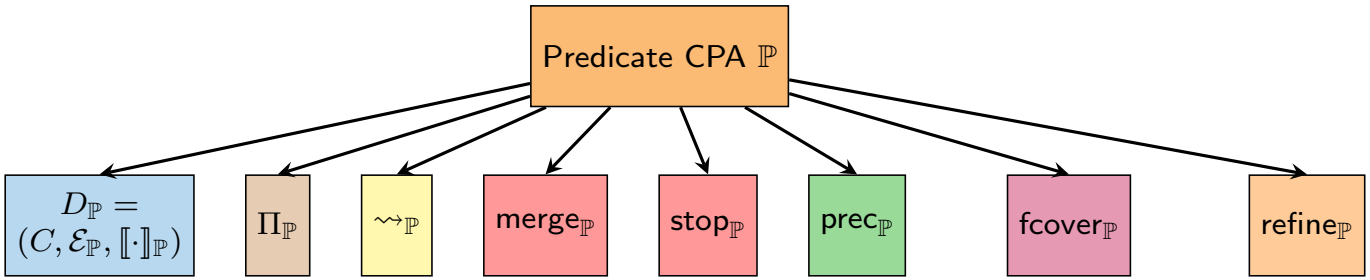
Ziele

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen
2. Verstehen der Unterschiede und Kernideen der Ansätze
3. Ermitteln des Potentials für Erweiterungen und Kombinationen
4. Solide Plattform für experimentelle Forschung

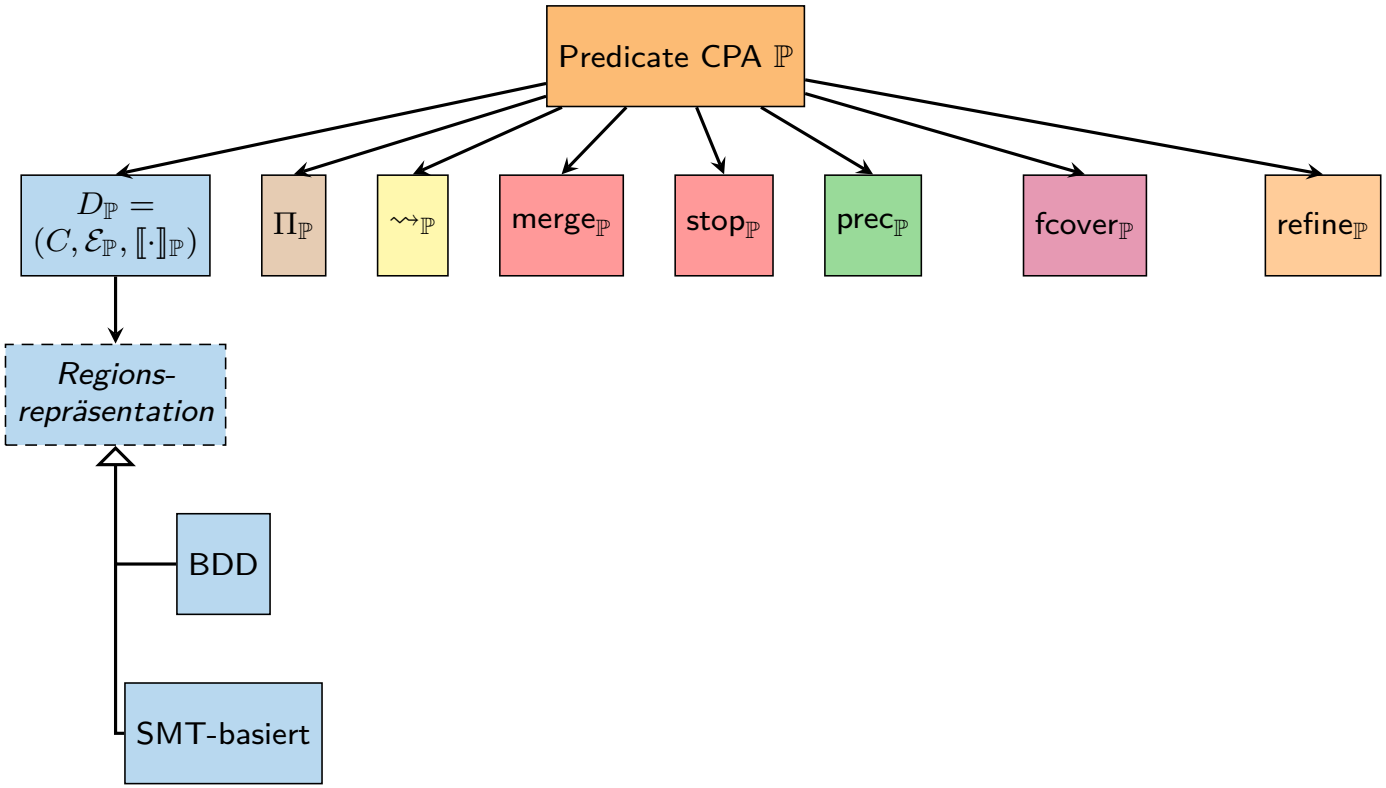
Ansatz

- Erforschen und, soweit nötig, Umformulieren der Algorithmen
- Entwurf eines konfigurierbaren Rahmenwerks für Prädikatenanalysen: **Predicate CPA**
- Ausdrücken vorhandener Algorithmen im gemeinsamen Rahmenwerk
  - Prädikatenabstraktion
  - IMPACT
  - Bounded Model-Checking
  - k-Induktion
- Implementierung des Rahmenwerks (in CPACHECKER)

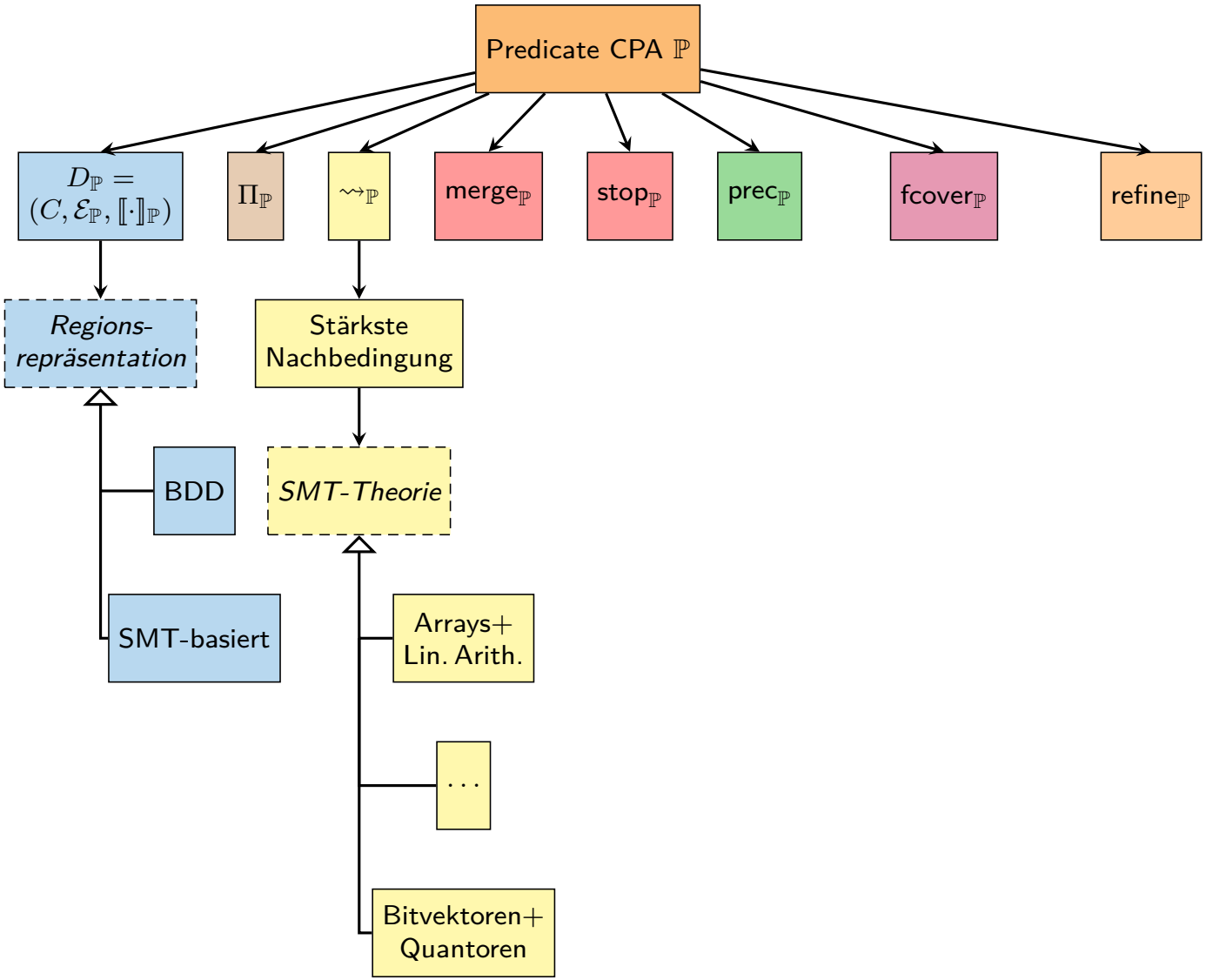
Predicate CPA



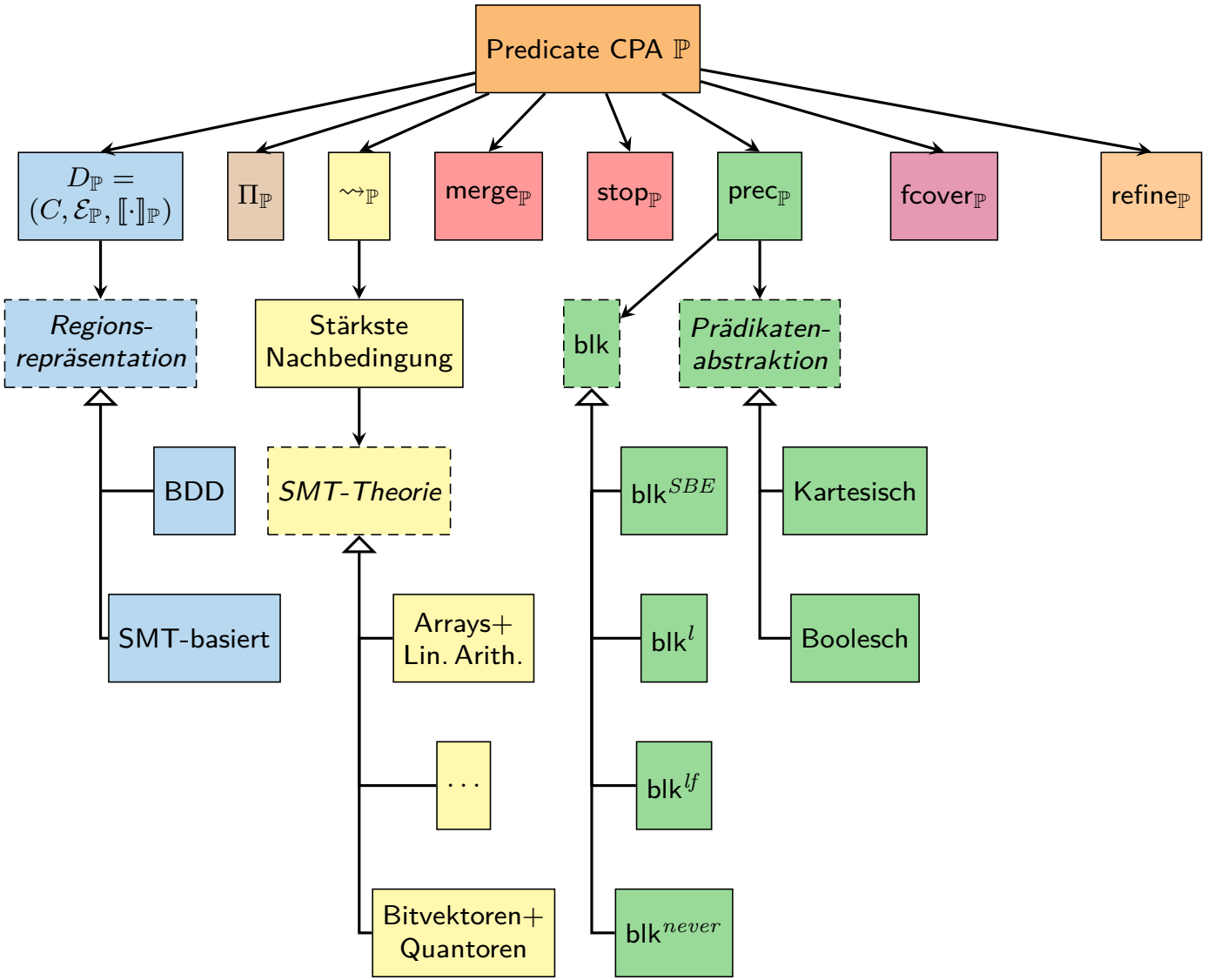
Predicate CPA



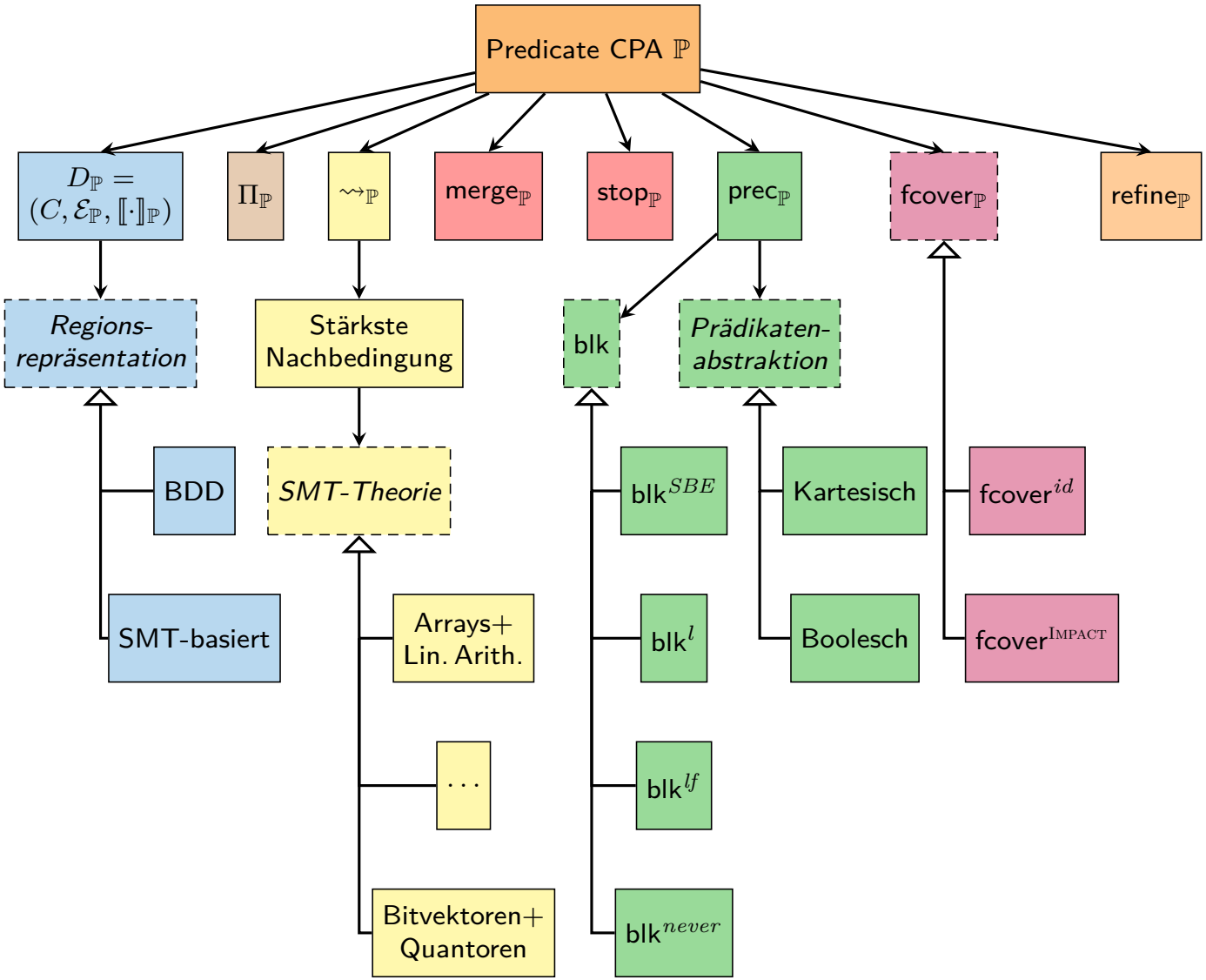
Predicate CPA



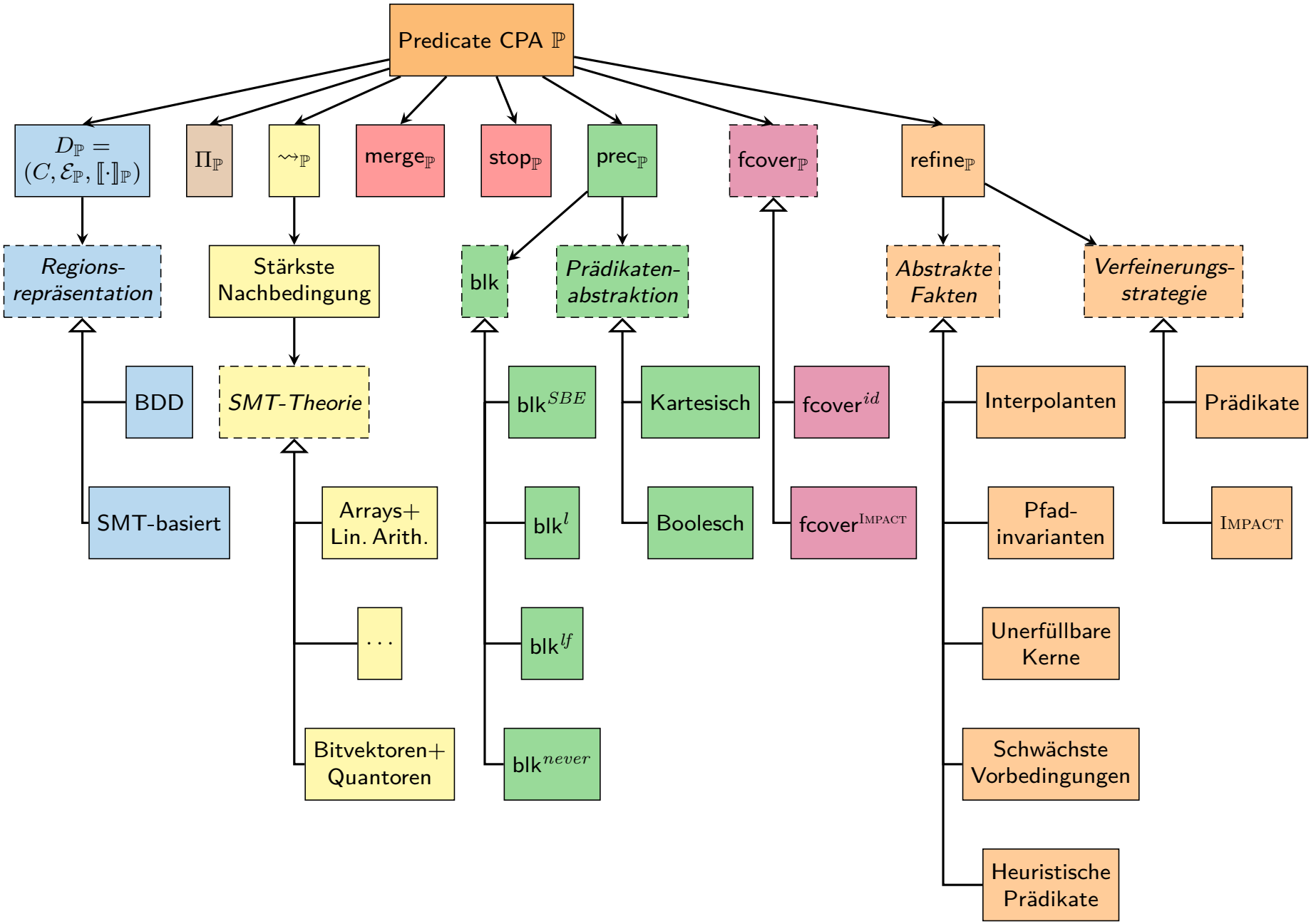
Predicate CPA



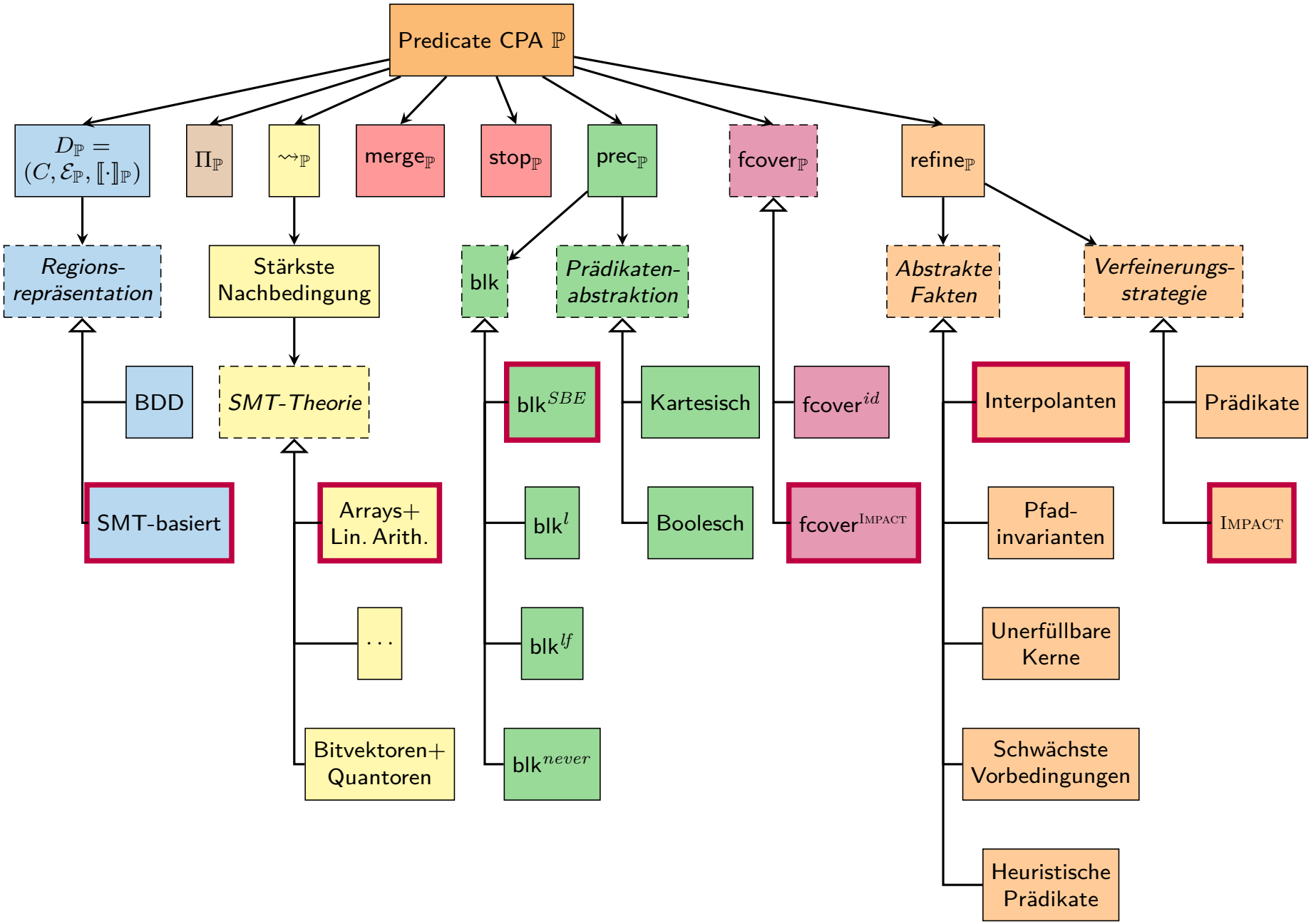
Predicate CPA



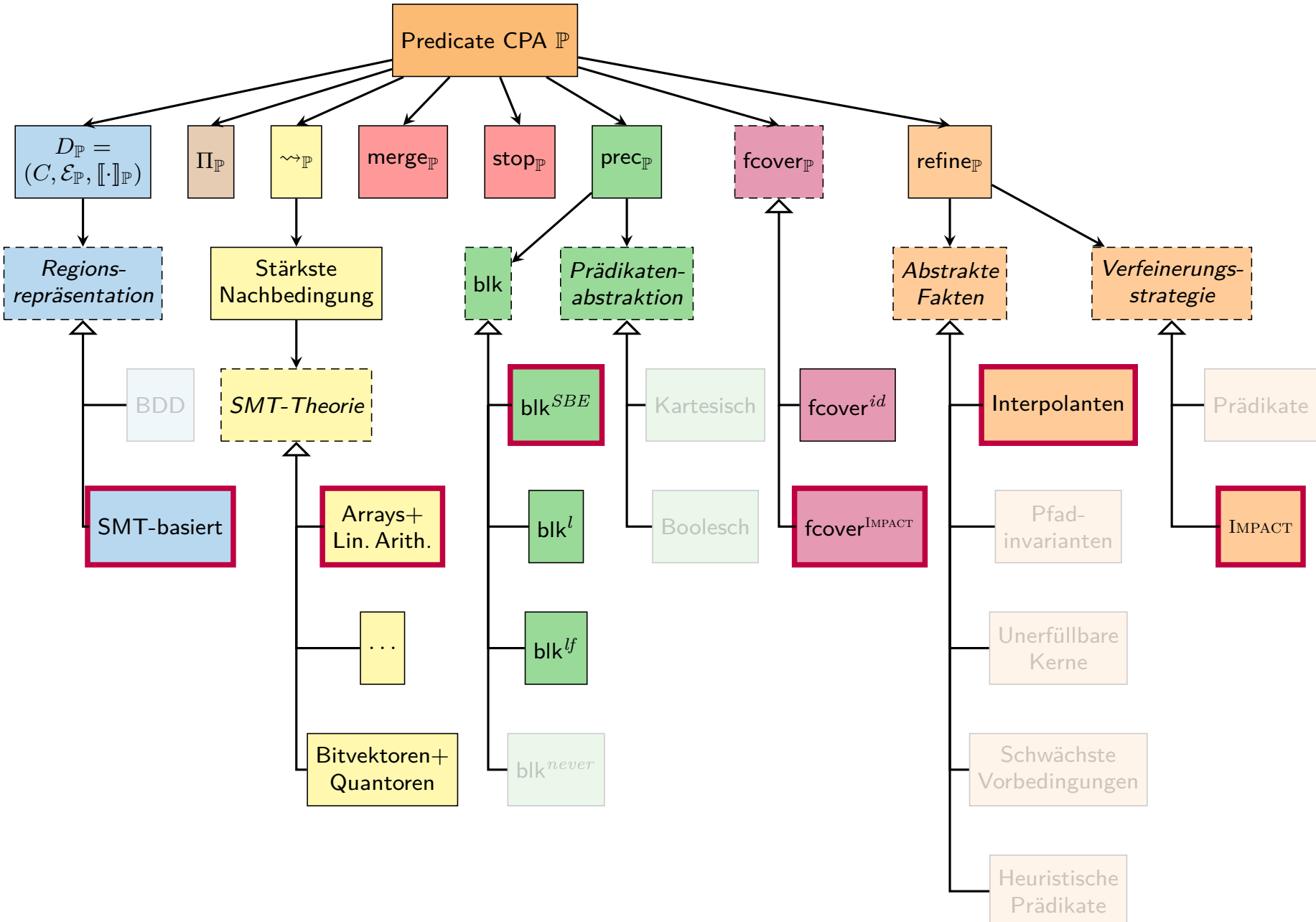
Predicate CPA



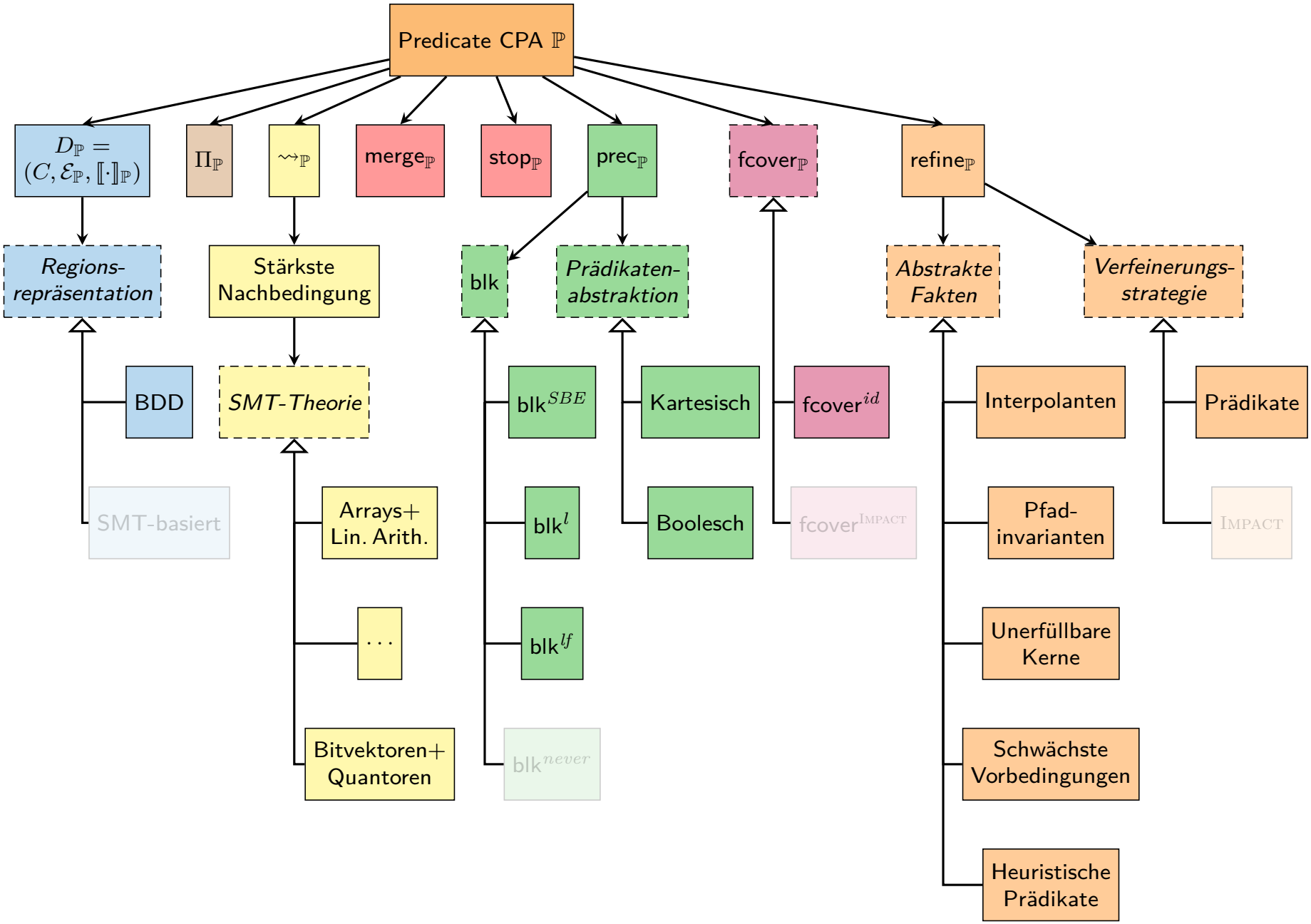
Predicate CPA für IMPACT



Predicate CPA für IMPACT



Predicate CPA für Prädikatenabstraktion



Evaluation: Nutzen des Rahmenwerks

- 4 vorhandene Ansätze erfolgreich integriert
- Laufende Projekte zur Integration weiterer Ansätze
- Frühere Ergebnisse reproduziert
- Interessante Erkenntnisse über diese Ansätze gelernt
- Hohe Konfigurierbarkeit erlaubt neue Kombinationen und hybride Ansätze

Evaluation: Nutzen des Rahmenwerks

- 4 vorhandene Ansätze erfolgreich integriert
- Laufende Projekte zur Integration weiterer Ansätze
- Frühere Ergebnisse reproduziert
- Interessante Erkenntnisse über diese Ansätze gelernt
- Hohe Konfigurierbarkeit erlaubt neue Kombinationen und hybride Ansätze
- Bereits verwendet in erfolgreichen Forschungsprojekten von anderen Forschern, z. B.
  - Block-abstraction memoization [ICFEM'12]
  - Refinement selection [SPIN'15]
  - Local policy iteration [VMCAI'16]
  - ...

Evaluation: Nutzen der Implementierung

- Eingesetzt in anderen Forschungsprojekten:
  - Conditional Model-Checking [FSE'12]
  - Verifikation rekursiver Programme [SAS'14]
  - Verifikationszeugen [FSE'15, FSE'16]

Evaluation: Nutzen der Implementierung

- Eingesetzt in anderen Forschungsprojekten:
  - Conditional Model-Checking [FSE'12]
  - Verifikation rekursiver Programme [SAS'14]
  - Verifikationszeugen [FSE'15, FSE'16]
- Ermöglicht experimentelle Studien:
  - SMT-based software model checking:
    An experimental comparison of four algorithms
    [VSTTE'16]
  - Vergleich von SMT-Solvern und Theorien
    - 120 verschiedene Konfigurationen auf 5 594 Programmen
    - Wichtige Erkenntnisse darüber wie SMT-Solver und -Theorien Benchmark-Ergebnisse und Folgerungen beeinflussen können

Beide Studien vorher nicht möglich!

Evaluation: Vergleich mit Stand der Technik

- Stand der Technik sichtbar in der Intl. Competition on Software Verification (SV-COMP)
- Implementierung gewann 4 Medaillen im ersten Jahr (SV-COMP'12)
- Trug bei zu 48 weiteren Medaillen
- Kurt-Gödel-Medaille der Kurt-Gödel-Gesellschaft verliehen

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen
2. Verstehen der Unterschiede und Kernideen der Ansätze
3. Ermitteln des Potentials für Erweiterungen und Kombinationen
4. Solide Plattform für experimentelle Forschung

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen ✓
   - Formal definiert und für 4 Analysen eingesetzt:
     publiziert im Journal of Automated Reasoning (JAR)
   - Übernommen durch Andere für weitere Analysen
2. Verstehen der Unterschiede und Kernideen der Ansätze
3. Ermitteln des Potentials für Erweiterungen
   und Kombinationen
4. Solide Plattform für experimentelle Forschung

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen ✓
   - Formal definiert und für 4 Analysen eingesetzt:
     publiziert im Journal of Automated Reasoning (JAR)
   - Übernommen durch Andere für weitere Analysen
2. Verstehen der Unterschiede und Kernideen der Ansätze ✓
   - Interessante Erkenntnisse gefunden
3. Ermitteln des Potentials für Erweiterungen
   und Kombinationen
4. Solide Plattform für experimentelle Forschung

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen ✓
   - Formal definiert und für 4 Analysen eingesetzt: publiziert im Journal of Automated Reasoning (JAR)
   - Übernommen durch Andere für weitere Analysen
2. Verstehen der Unterschiede und Kernideen der Ansätze ✓
   - Interessante Erkenntnisse gefunden
3. Ermitteln des Potentials für Erweiterungen und Kombinationen ✓
   - Neue Kombinationen von Features und Algorithmen schon verfügbar
   - Großes Potential für zukünftige Forschung
4. Solide Plattform für experimentelle Forschung

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen ✓
   - Formal definiert und für 4 Analysen eingesetzt: publiziert im Journal of Automated Reasoning (JAR)
   - Übernommen durch Andere für weitere Analysen
2. Verstehen der Unterschiede und Kernideen der Ansätze ✓
   - Interessante Erkenntnisse gefunden
3. Ermitteln des Potentials für Erweiterungen und Kombinationen ✓
   - Neue Kombinationen von Features und Algorithmen schon verfügbar
   - Großes Potential für zukünftige Forschung
4. Solide Plattform für experimentelle Forschung ✓
   - Hochrangige Implementierung (open source)
   - Eingesetzt für mehrere experimentelle Studien

Zusammenfassung

1. Vereinheitlichendes Rahmenwerk für Prädikatenanalysen ✓
   - Formal definiert und für 4 Analysen eingesetzt: publiziert im Journal of Automated Reasoning (JAR)
   - Übernommen durch Andere für weitere Analysen
2. Verstehen der Unterschiede und Kernideen der Ansätze ✓
   - Interessante Erkenntnisse gefunden
3. Ermitteln des Potentials für Erweiterungen und Kombinationen ✓
   - Neue Kombinationen von Features und Algorithmen schon verfügbar
   - Großes Potential für zukünftige Forschung
4. Solide Plattform für experimentelle Forschung ✓
   - Hochrangige Implementierung (open source)
   - Eingesetzt für mehrere experimentelle Studien