CSC 548 - Artificial Intelligence II, Spring 2019

Resolution Theorem Proving

Proof Methods

- Proof methods divide into (roughly) two kinds:
- Application of inference rules
  - Legitimate (sound) generation of new sentences from old.
  - Proof $=$ a sequence of inference rule applications.
  - Typically require translation of sentences into a normal form.
  - Meaning is context-independent
- Model checking
  - Truth table enumeration
  - Improved backtracking
  - Heuristic search in model space (sound but not complete)

Propositional Resolution

- Resolution rule:

$$\frac{\begin{array}{c}\alpha \vee \beta \\ \neg\beta \vee \gamma\end{array}}{\alpha \vee \gamma}$$

- Resolution refutation:
  - Convert all sentences to conjunctive normal form (CNF)
  - Negate the desired conclusion (converted to CNF)
  - Apply resolution rule until either
    - derive false (a contradiction)
    - cannot apply any more
- Resolution refutation is sound and complete

Conjunctive Normal Form (CNF)

- Conjunctive Normal Form (CNF):
  conjunction of disjunction of literals
- Example: $(A \vee \neg B) \wedge (B \vee \neg C \vee \neg D)$
- $(A \vee \neg B)$ is a clause
- A and $\neg B$ are literals, each of which is a variable or a negation of a variable
- Every sentence in propositional logic can be written in CNF

Conversion to CNF

1. Eliminate $\leftrightarrow$ replacing $\alpha \leftrightarrow \beta$ with $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$
2. Eliminate $\rightarrow$, replacing $\alpha \rightarrow \beta$ with $\neg\alpha \vee \beta$.
3. Move $\neg$ inwards using de Morgan's rules and double-negation
4. Apply distributivity law ($\vee$ over $\wedge$) and flatten

- Example:

$P \leftrightarrow (Q \vee R)$
$(P \rightarrow (Q \vee R)) \wedge ((Q \vee R) \rightarrow P)$
$(\neg P \vee Q \vee R) \wedge (\neg(Q \vee R) \vee P)$
$(\neg P \vee Q \vee R) \wedge ((\neg Q \wedge \neg R) \vee P)$
$(\neg P \vee Q \vee R) \wedge (\neg Q \vee P) \wedge (\neg R \vee P)$

Resolution Example

Prove R

1	$P \vee Q$
2	$P \rightarrow R$
3	$Q \rightarrow R$

Step	Formula	Derivation
1	$P \vee Q$	Given
2	$\neg P \vee R$	Given
3	$\neg Q \vee R$	Given
4	$\neg R$	Negated Conclusion
5	$Q \vee R$	1,2
6	$\neg P$	2,4 $*$
7	$\neg Q$	3,4
8	R	5,7
9	$\Box$	4,8

The Power of False

Prove Z

1	P
2	$\neg P$

Step	Formula	Derivation
1	P	Given
2	$\neg P$	Given
4	$\neg Z$	Negated Conclusion
5	$\square$	1,2

- Note that $(P \vee \neg P) \rightarrow Z$ is valid
- Any conclusion follows from a contradiction – and so strict logic systems are brittle.

Proof Strategies

- Unit preference: prefer a resolution step involving an unit clause (clause with one literal).
  - Produces shorter clause – which is good since we are trying to produce a zero-length clause, that is, a contradiction.
- Set of support: choose a resolution involving the negated goal or any clause derived from the negated goal.
  - We are trying to produce a contradiction that follows from the negated goal, so these are "relevant" clauses.
  - If a contradiction exists, one can find one using the set of support strategy.

First-Order Resolution

- Syllogism

$$\frac{\begin{array}{c}\forall x\ P(x) \rightarrow Q(x)\\ P(A)\end{array}}{Q(A)}$$

- Equivalent by definition of implication

$$\frac{\begin{array}{c}\forall x\ \neg P(x) \vee Q(x)\\ P(A)\end{array}}{Q(A)}$$

- Substitute A for x, then propositional resolution

$$\frac{\begin{array}{c}\forall x\ \neg P(x) \vee Q(x)\\ P(A)\end{array}}{Q(A)}$$

- The key is finding the correct substitutions for the variables.

Substitutions

- An atomic sentence: $P(x, F(y), B)$

Substitution Instances	Substitution $\{v_1/t_1, \ldots, v_n/t_n\}$	Comment
$P(z, F(w), B)$	$\{x/z, y/w\}$	alphabetic variant
$P(x, F(A), B)$	$\{y/A\}$	
$P(G(z), F(A), B)$	$\{x/G(z), y/A\}$	
$P(C, F(A), B)$	$\{x/C, y/A\}$	ground instance

Unification

- Expressions ω_1 and ω_2 are *unifiable* iff there exists a substitution σ such that $\omega_1\sigma = \omega_2\sigma$
- Let $\omega_1 = x$ and $\omega_2 = y$, the following are unifiers

σ	$\omega_1\sigma$	$\omega_2\sigma$
$\{y/x\}$	x	x
$\{x/y\}$	y	y
$\{x/F(A), y/F(A)\}$	$F(A)$	$F(A)$
$\{x/A, y/A\}$	A	A

Most General Unifier

- g is the *most general unifier* of ω_1 and ω_2 iff for all unifiers σ, there exists σ' such that $\omega_1\sigma = (\omega_1 g)\sigma'$ and $\omega_2\sigma = (\omega_2 g)\sigma'$

ω_1	ω_2	MGU
$P(x)$	$P(A)$	$\{x/A\}$
$P(F(x), y, G(x)$	$P(F(x), x, G(x)$	$\{y/x\}$ or $\{x/y\}$
$P(F(x), y, G(y)$	$P(F(x), z, G(x)$	$\{y/x, z/x\}$
$P(x, B, B)$	$P(A, y, z)$	$\{x/A, y/B, z/B\}$
$P(G(F(v)), G(u))$	$P(x, x)$	$\{x/G(F(v)), u/F(v)\}$
$P(x, F(x))$	$P(x, x)$	None

Inference Using Unification

- Inference rule:

$$\frac{\begin{array}{c}\forall x\ \neg P(x) \vee Q(x)\\ P(A)\end{array}}{Q(A)}$$

- For universally quantified variables, find MGU $\{x/A\}$ and proceed as in propositional resolution.

Resolution with Variables

- First-order resolution rule:

$$\frac{\begin{array}{c}\alpha \vee \phi \\ \neg\psi \vee \beta\end{array}}{(\alpha \vee \beta)\sigma}$$

  where $\sigma = \mathit{MGU}(\psi, \phi)$

- Example

$$\frac{\begin{array}{c}P(x) \vee Q(x, y) \\ \neg P(A) \vee R(B, z)\end{array}}{(Q(x, y) \vee R(B, z))\sigma}$$

  $\sigma = \{x/A\}$

Resolution with Variables Example

- Another example:

$$\frac{\begin{array}{c} P(x) \vee Q(x,y) \\ \neg P(A) \vee R(B,x) \end{array}}{(Q(x,y) \vee R(B,x))\sigma}$$

- All variables are implicitly universally quantified and the scope of a variable is local to a clause. Need to rename to keep variables distinct.

$$\frac{\begin{array}{c} \forall x_1 y\ P(x_1) \vee Q(x_1,y) \\ \forall x_2\ \neg P(A) \vee R(B,x_2) \end{array}}{(Q(x_1,y) \vee R(B,x_2))\sigma}$$

$\sigma = \{x_1/A\}$

Resolution

- Input are sentences in conjunctive normal form with no apparent quantifiers (implicit universal quantifiers).
- How to we go from the full range of sentences in FOL, with the full range of quantifiers, to sentences that enable us to use resolution as our single inference rule?
- We will convert the input sentences into a new normal form called *clausal form* (also called prenex normal form).

Converting to Clausal Form

1. Eliminate implications ($\rightarrow$ and $\leftrightarrow$)
2. Drive in negation (deMorgan's laws and quantifiers)
3. Rename variables apart
4. Skolemize
   - substitute a brand new name for each existentially quantified variable
   - substitute a new function of all universally quantified variables in enclosing scopes for each existentially quantified variable
5. Drop universal quantifiers
6. Convert to CNF
7. Rename the variables in each clause

Example: Convert to Clausal Form

- a. "John owns a dog"

$\exists x\ Dog(x) \land Owns(John, x)$
$Dog(Fido) \land Owns(John, Fido)$

- b. "Anyone who owns a dog is a lover of animals"

$\forall x\ (\exists y\ Dog(y) \land Owns(x, y)) \rightarrow LovesAnimals(x)$
$\forall x\ (\neg\exists y\ Dog(y) \land Owns(x, y)) \lor LovesAnimals(x)$
$\forall x\ \forall y\ \neg(Dog(y) \land Owns(x, y)) \lor LovesAnimals(x)$
$\forall x\ \forall y\ \neg Dog(y) \lor \neg Owns(x, y) \lor LovesAnimals(x)$
$\neg Dog(y) \lor \neg Owns(x, y) \lor LovesAnimals(x)$

Example: Convert to Clausal Form

- c. "Lovers of animals do not kill animals"

$\forall x \; LovesAnimals(x) \rightarrow (\forall y \; Animal(y) \rightarrow \neg Kill(x, y))$
$\forall x \; \neg LovesAnimals(x) \vee (\forall y \; Animal(y) \rightarrow \neg Kill(x, y))$
$\forall x \; \neg LovesAnimals(x) \vee (\forall y \; \neg Animal(y) \vee \neg Kill(x, y))$
$\neg LovesAnimals(x) \vee \neg Animal(y) \vee \neg Kill(x, y)$

- d. "Either John killed Tuna or curiosity killed Tuna"

$$\boxed{Kill(John, Tuna) \vee Kill(Curiosity, Tuna)}$$

- e. "Tuna is a cat"

$$\boxed{Cat(Tuna)}$$

- f. "All cats are animals"

$$\boxed{\neg Cat(x) \vee Animal(x)}$$

Example: Resoution Proof – Curiosity Killed the Cat

1	$Dog(Fido)$	a
2	$Owns(John, Fido)$	a
3	$\neg Dog(y) \vee \neg Owns(x, y) \vee LovesAnimals(x)$	b
4	$\neg LovesAnimals(x) \vee \neg Animal(y) \vee \neg Kill(x, y)$	c
5	$Kill(John, Tuna) \vee Kill(Curiosity, Tuna)$	d
6	$Cat(Tuna)$	e
7	$\neg Cat(x) \vee Animal(x)$	f
8	$\neg Kill(Curiostiy, Tuna)$	Neg
9	$Kill(John, Tuna)$	5,8
10	$Animal(Tuna)$	6,7 $\{x/Tuna\}$
11	$\neg LovesAnimals(John) \vee \neg Animal(Tuna)$	4,9 $\{x/John, y/Tuna\}$
12	$\neg LovesAnimals(John)$	10,11
13	$Dog(y) \vee Owns(John, y)$	3,12 $\{x/John\}$
14	$\neg Dog(Fido)$	13,2 $\{x/Fido\}$
15	$\square$	14,1