Automatic Reasoning: SAT solving on CNFs suffices

Albert Oliveras and Enric Rodríguez-Carbonell and Javier Larrosa

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Overview of the session

- Reduction to SAT
- Conversion to CNF
  - Via truth table
  - Via distributivity
  - Tseitsin

Recall: Usual Queries

Let F and G be arbitrary formulas. Then:

- F is satisfiable (also consistent) if it has at least one model
- F is unsatisfiable (also a contradiction) if it has no model
- F is a tautology if every interpretation is a model of F
- G is a logical consequence of F (F entails G), denoted $F \models G$, if every model of F is a model of G
- F and G are logically equivalent, denoted $F \equiv G$, if F and G have the same models

Reduction to SAT

Assume we have a black-box **SAT** that given a formula F:

- **SAT**(F)=YES iff F is satisfiable
- **SAT**(F)=NO iff F is unsatisfiable

How to reuse SAT for detecting tautology, logical consequences, ...?

- F tautology iff **SAT**($\neg F$)=NO
- $F \models G$ iff **SAT**($F \wedge \neg G$)=NO
- $F \equiv G$ iff **SAT**($(F \wedge \neg G) \vee (\neg F \wedge G)$)=NO

Hence, a single tool suffices.

A VERY USEFUL TOOL: black-box SAT

Complexity of SAT

SAT is an NP-Complete problem

- All known algorithms have worst-case exponential cost on the input size
- It is exponentially expensive solving the problem, but only polynomially expensive checking if an assignment is a model
- Real problems may not be worst-case instances for some algorithms

CNF

In order to construct our **SAT** black-box it would simplify our job to assume that the formula F has a given format.

- A literal is a prop. variable (p) or a negation of one ($\neg p$)
- A clause is a disjunction of zero or more literals ($l_1 \vee \ldots l_n$)
  - The empty clause (zero lits.) is denoted $\square$ and is unsatisfiable
- A formula is in Conjunctive Normal Form (CNF) if it is a conjunction of clauses
  - The empty formula (zero clauses) is satisfiable
- Tautology clauses (e.g. $r \vee p \vee \neg r$) are removed
- repeated literals (e.g. $p \vee p \vee \neg r$) are removed

CNF

Example:

$$p \wedge (q \vee \neg r) \wedge (q \vee p \vee \neg r) \text{ is in CNF}$$

Property: Every formula can be transformed into CNF

Transformation to CNF via truth table

Let us take the formula $F := (p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$
Its truth table is:

p	q	r	
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

- It is easy to compute a CNF for F:
  $(\ p \vee q \vee r) \quad \wedge \quad (p \vee \neg q \vee r) \quad \wedge$
  $(p \vee \neg q \vee \neg r)$
- This method may produce unnecessaryly large CNF formulas (e.g. $p \wedge q \wedge r$)

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow (p \wedge q) \vee (p \vee (\neg q \wedge r))$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow (p \wedge q) \vee (p \vee (\neg q \wedge r))$
2. $(p \wedge q) \vee (p \vee (\neg q \wedge r)) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow (p \wedge q) \vee (p \vee (\neg q \wedge r))$
2. $(p \wedge q) \vee (p \vee (\neg q \wedge r)) \Rightarrow (p \vee p \vee (\neg q \wedge r)) \wedge (q \vee p \vee (\neg q \wedge r)) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow (p \wedge q) \vee (p \vee (\neg q \wedge r))$
2. $(p \wedge q) \vee (p \vee (\neg q \wedge r)) \Rightarrow (p \vee p \vee (\neg q \wedge r)) \wedge (q \vee p \vee (\neg q \wedge r)) \Rightarrow$
   $(p \vee p \vee \neg q) \wedge (p \vee p \vee r) \wedge (q \vee p \vee \neg q) \wedge (q \vee p \vee r) \Rightarrow$

Tranformation to CNF via distributivity

1. Apply the three transformation rules up to completion:
   - $\neg\neg F \Rightarrow F$
   - $\neg(F \wedge G) \Rightarrow \neg F \vee \neg G$
   - $\neg(F \vee G) \Rightarrow \neg F \wedge \neg G$

   After that, the formula is in Negation Normal Form (NNF)
2. Now apply the distributivity rule up to completion:
   - $F \vee (G \wedge H) \Rightarrow (F \vee G) \wedge (F \vee H)$
3. remove tautology clauses and repeated literals

EXAMPLE: let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

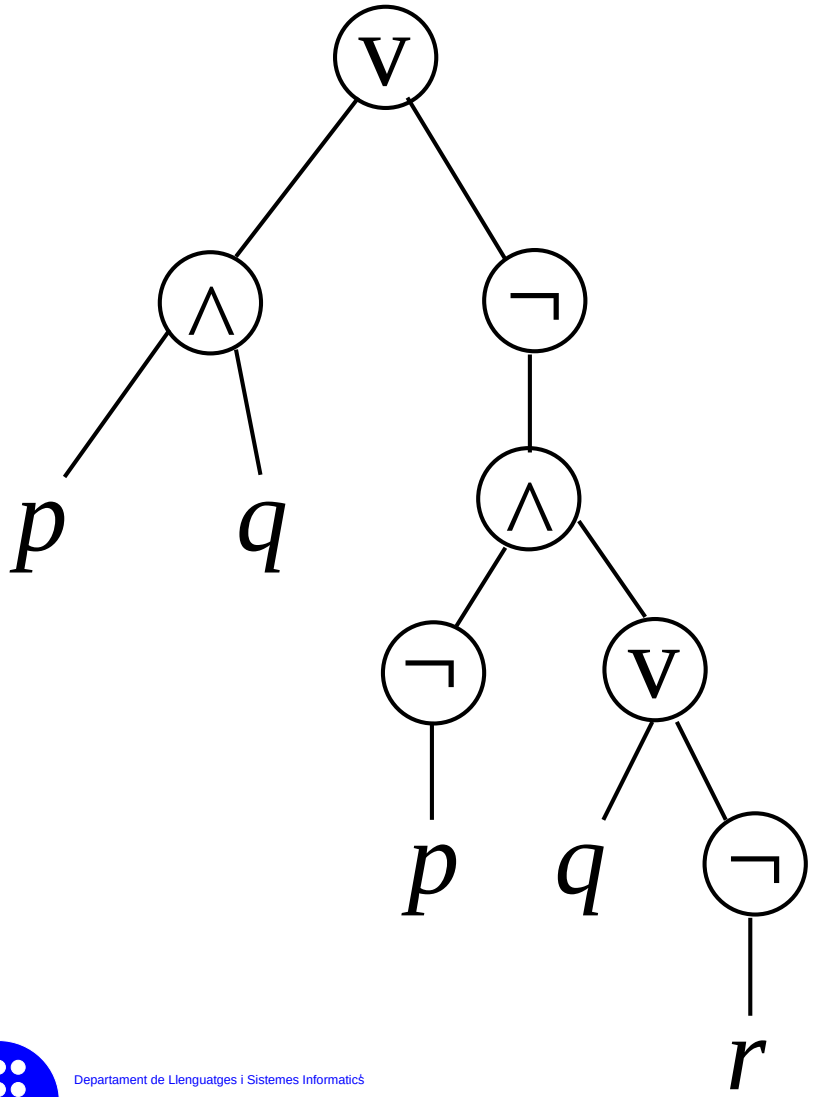
1. $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ ) \Rightarrow (p \wedge q) \vee (\neg\neg p \vee \neg(q \vee \neg r)) \Rightarrow$
   $(p \wedge q) \vee (p \vee (\neg q \wedge \neg\neg r)) \Rightarrow (p \wedge q) \vee (p \vee (\neg q \wedge r))$
2. $(p \wedge q) \vee (p \vee (\neg q \wedge r)) \Rightarrow (p \vee p \vee (\neg q \wedge r)) \wedge (q \vee p \vee (\neg q \wedge r)) \Rightarrow$
   $(p \vee p \vee \neg q) \wedge (p \vee p \vee r) \wedge (q \vee p \vee \neg q) \wedge (q \vee p \vee r) \Rightarrow$
   $(p \vee \neg q) \wedge (p \vee r) \wedge (q \vee p \vee r)$

Tranformation to CNF via distributivity

- This method may produce an exponential growth in the size of the formula
- **Example:**
  - $(p_0 \wedge p_1) \vee (p_2 \wedge p_3) =$
    $(p_0 \vee p_2) \wedge (p_0 \vee p_3) \wedge (p_1 \vee p_2) \wedge (p_1 \vee p_3)$
  - $(p_0 \wedge p_1) \vee (p_2 \wedge p_3) \vee (p_4 \wedge p_5) =$
    $(p_0 \vee p_2 \vee p_4) \wedge (p_0 \vee p_3 \vee p_4) \wedge (p_1 \vee p_2 \vee p_4) \wedge (p_1 \vee p_3 \vee p_4) \wedge$
    $(p_0 \vee p_2 \vee p_5) \wedge (p_0 \vee p_3 \vee p_5) \wedge (p_1 \vee p_2 \vee p_5) \wedge (p_1 \vee p_3 \vee p_5)$
  - ...
  - In general: $(\wedge_i D_i) \vee (p \vee q) = (\wedge_i (D_i \vee p)) \wedge (\wedge_i (D_i \vee q))$

Tranformation to CNF via Tseitin

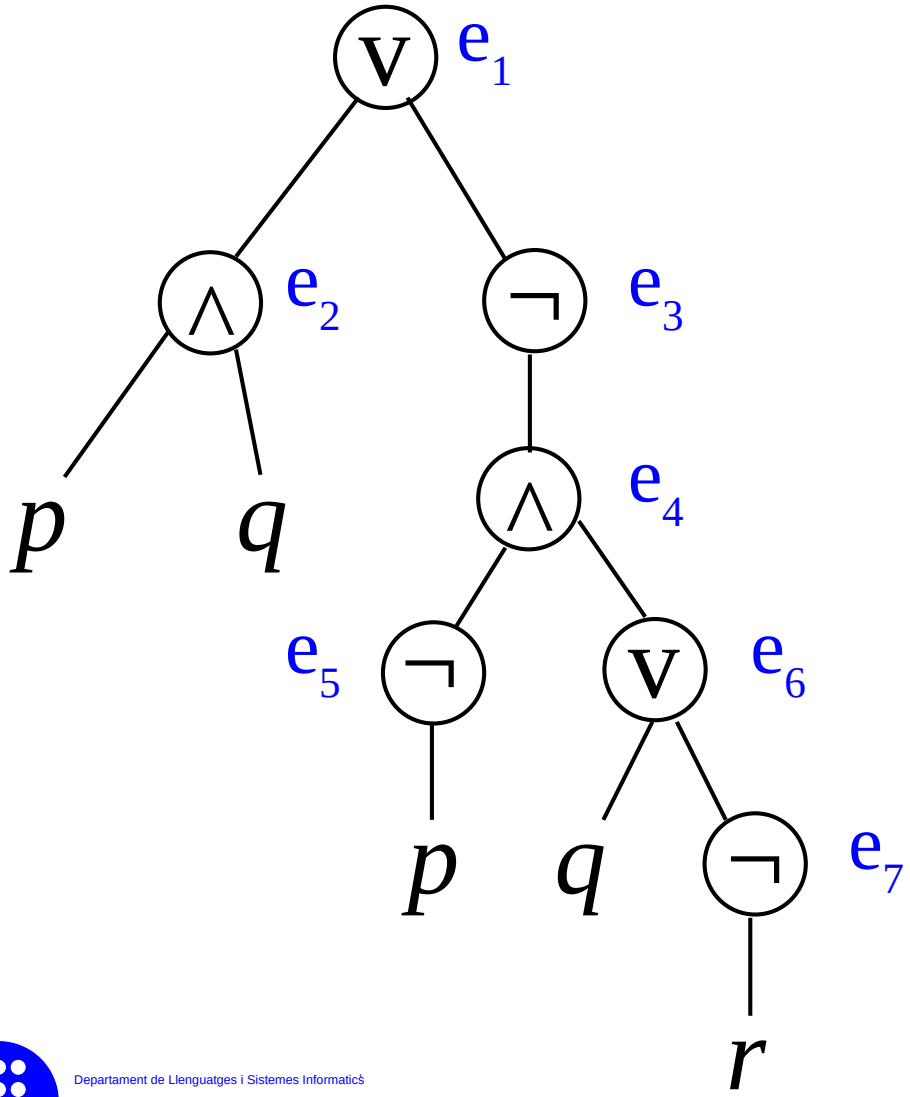
Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$



Departament de Llenguatges i Sistemes Informaticś
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin

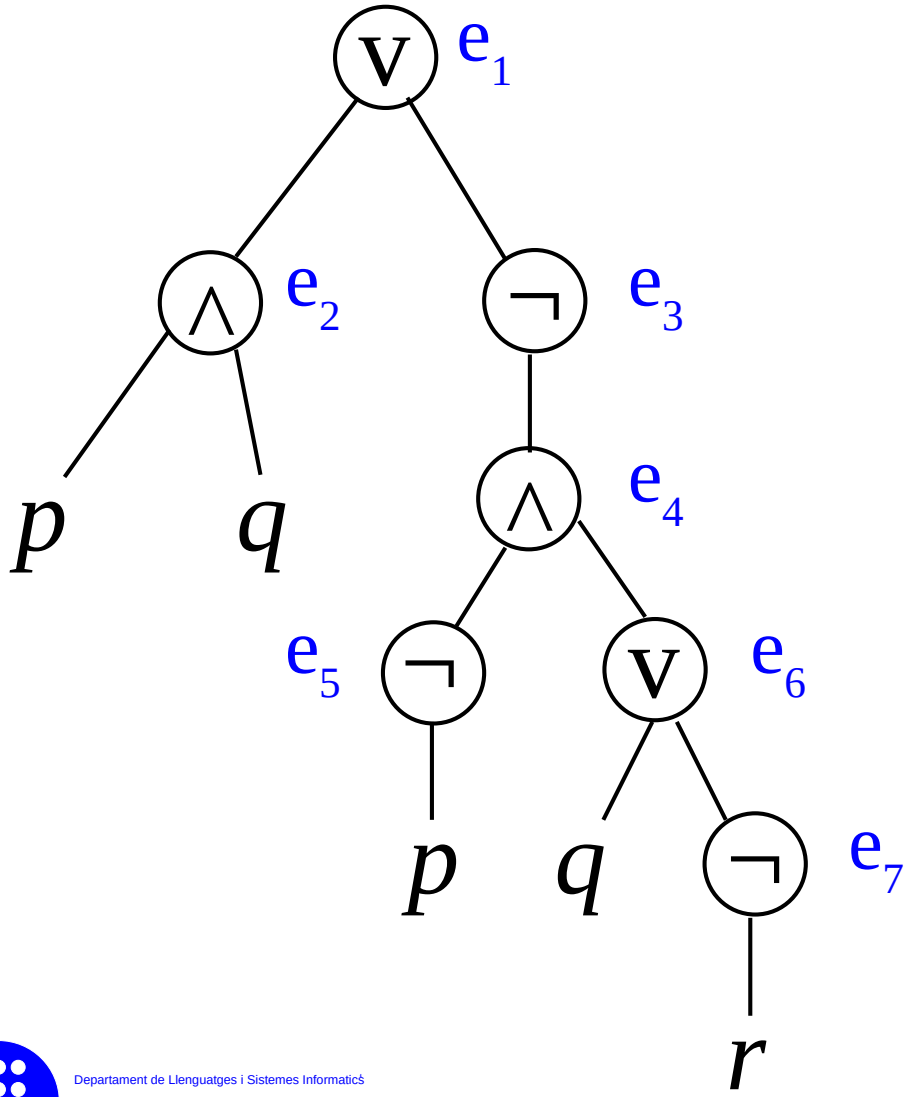
Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$



Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin

Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$

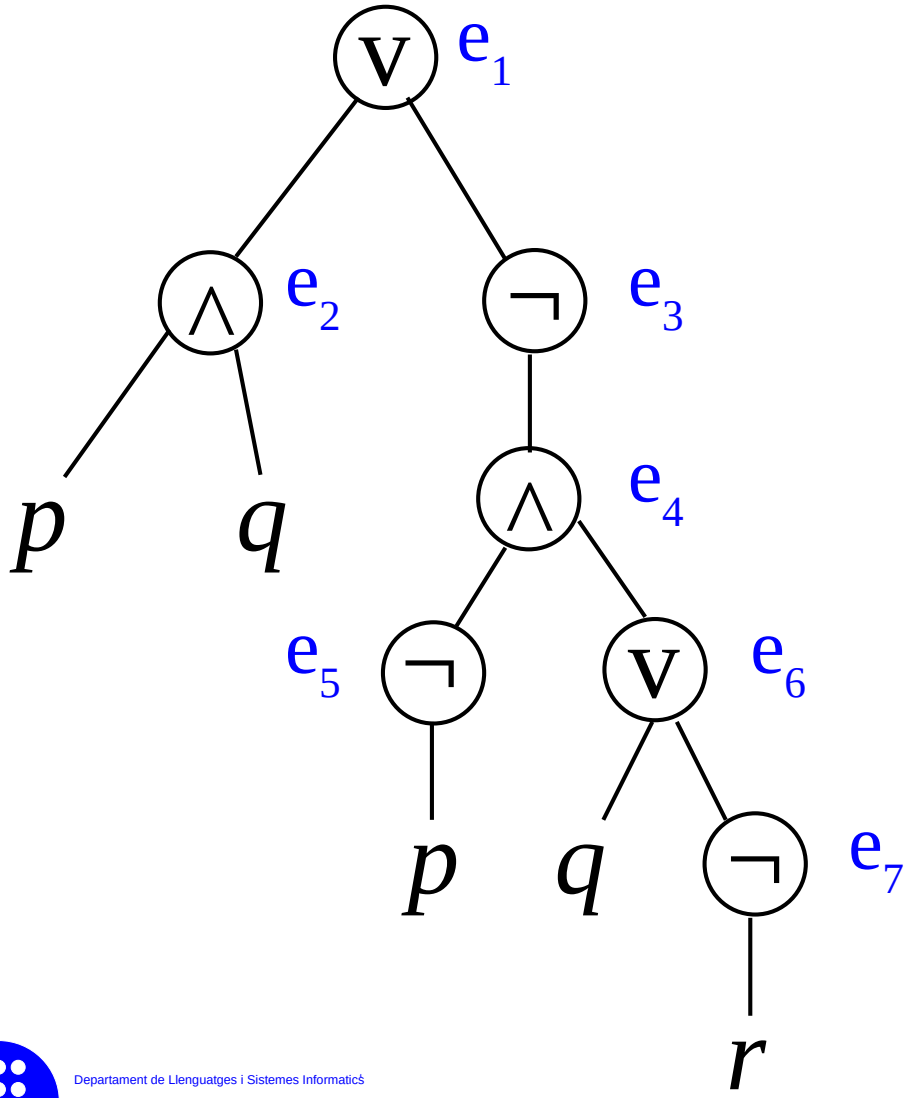


- e_1
- $e_1 \leftrightarrow e_2 \vee e_3$
- $e_2 \leftrightarrow p \wedge q$
- $e_3 \leftrightarrow \neg e_4$
- $e_4 \leftrightarrow e_5 \wedge e_6$
- $e_5 \leftrightarrow \neg p$
- $e_6 \leftrightarrow q \vee \neg e_7$
- $e_7 \leftrightarrow \neg r$

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin

Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$



- e_1
- $e_1 \leftrightarrow e_2 \vee e_3$

  $\neg e_1 \vee e_2 \vee e_3$

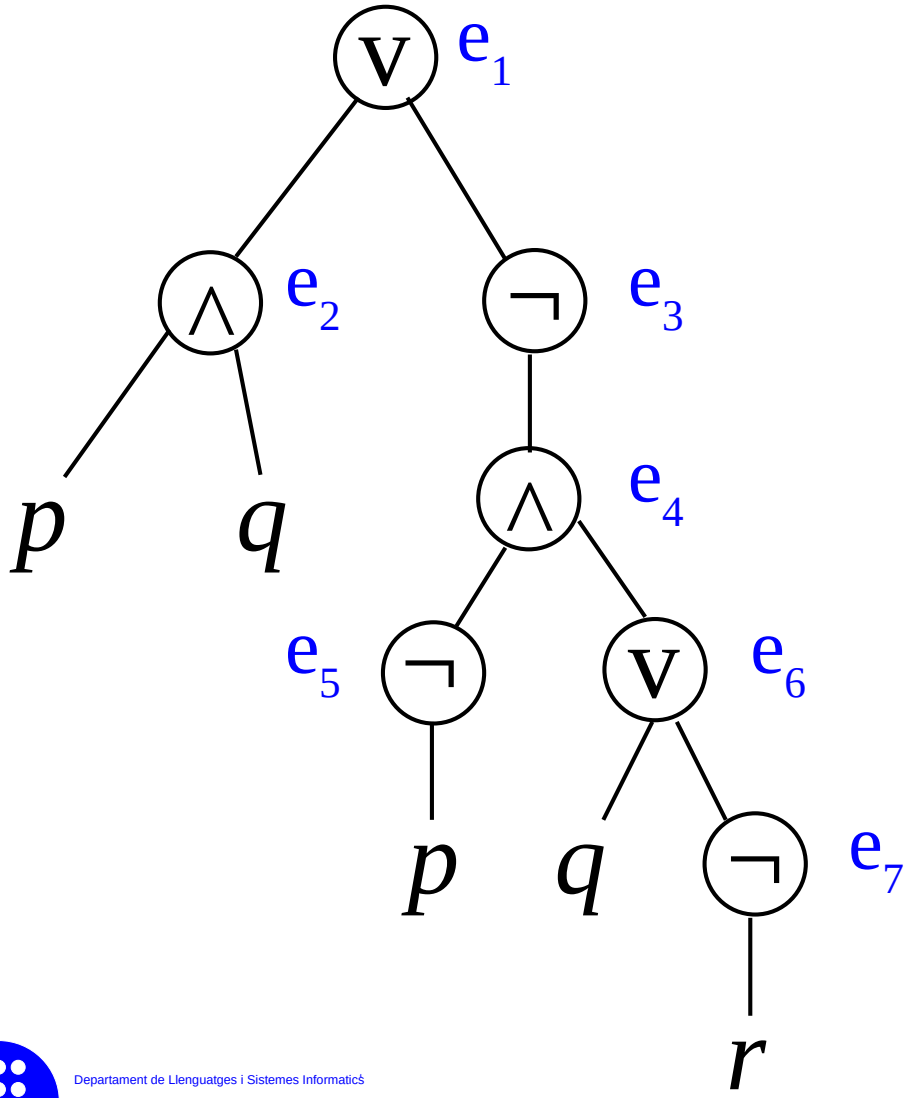
  $\neg e_2 \vee e_1$

  $\neg e_3 \vee e_1$
- $e_2 \leftrightarrow p \wedge q$
- $e_3 \leftrightarrow \neg e_4$
- $e_4 \leftrightarrow e_5 \wedge e_6$
- $e_5 \leftrightarrow \neg p$
- $e_6 \leftrightarrow q \vee \neg e_7$
- $e_7 \leftrightarrow \neg r$

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin

Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$



- e_1
- $e_1 \leftrightarrow e_2 \vee e_3$

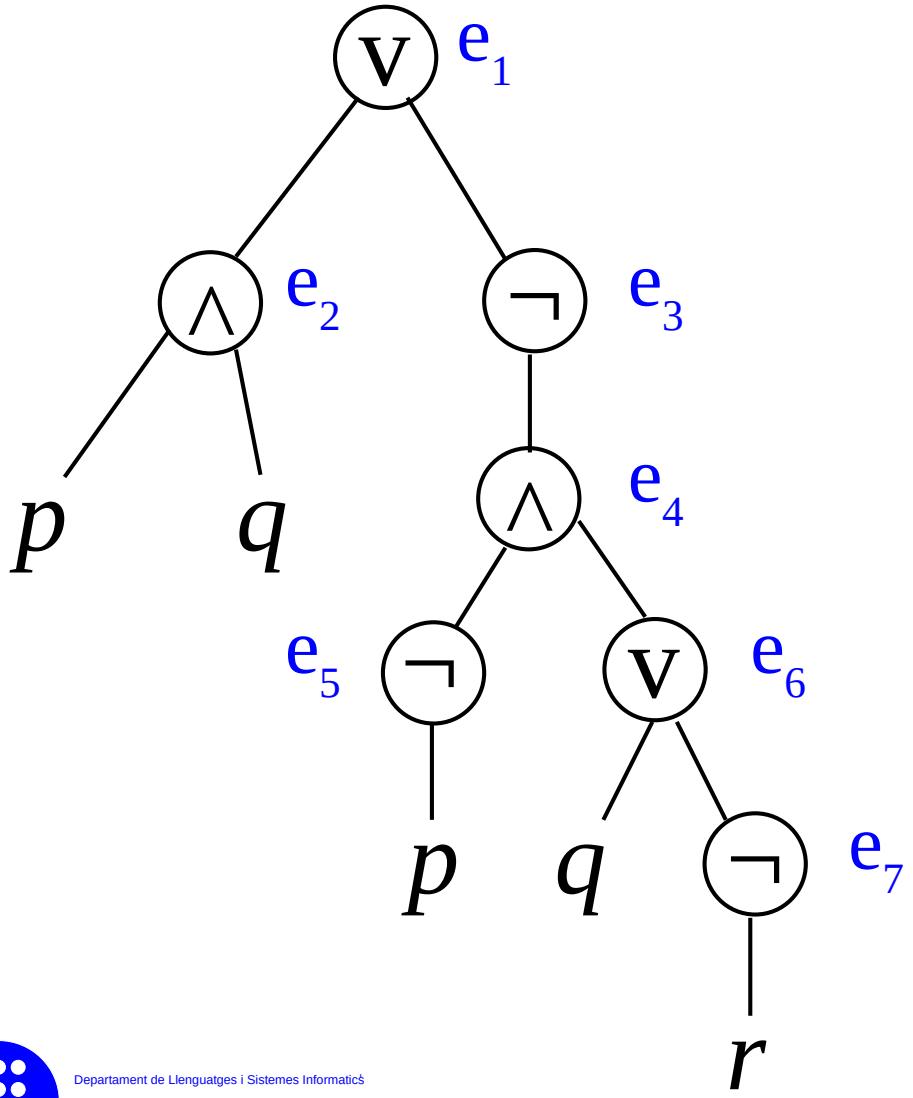
  $\neg e_1 \vee e_2 \vee e_3$
  $\neg e_2 \vee e_1$
  $\neg e_3 \vee e_1$
- $e_2 \leftrightarrow p \wedge q$

  $\neg p \vee \neg q \vee e_2$
  $\neg e_2 \vee p$
  $\neg e_2 \vee q$
- $e_3 \leftrightarrow \neg e_4$
- $e_4 \leftrightarrow e_5 \wedge e_6$
- $e_5 \leftrightarrow \neg p$
- $e_6 \leftrightarrow q \vee \neg e_7$
- $e_7 \leftrightarrow \neg r$

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin

Let F be $(p \wedge q) \vee \neg(\ \neg p \wedge (q \vee \neg r)\ )$



- e_1
- $e_1 \leftrightarrow e_2 \vee e_3$

  $\neg e_1 \vee e_2 \vee e_3$
  $\neg e_2 \vee e_1$
  $\neg e_3 \vee e_1$
- $e_2 \leftrightarrow p \wedge q$

  $\neg p \vee \neg q \vee e_2$
  $\neg e_2 \vee p$
  $\neg e_2 \vee q$
- $e_3 \leftrightarrow \neg e_4$

  $\neg e_3 \vee \neg e_4$
  $e_3 \vee e_4$
- $e_4 \leftrightarrow e_5 \wedge e_6$
- $e_5 \leftrightarrow \neg p$
- $e_6 \leftrightarrow q \vee \neg e_7$
- $e_7 \leftrightarrow \neg r$

Departament de Llenguatges i Sistemes Informatics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Tranformation to CNF via Tseitin (2)

- Variations of Tseitin are the ones used in practice
- Tseitin does not produce an equivalent CNF (the problem is in the *auxiliary* variables)
- Given F, the CNF obtained has three important properties:
  - It is equisatisfiable to F
  - Any model of CNF can be projected to the variables in F giving a model of F
  - Any model of F can be completed to a model of the CNF
- Hence no model is lost nor added in the conversion

Transformation to CNF via Tseitin (3)

- Tseitin does not produce an equivalent CNF
- $F = p \vee (q \wedge r)$
- $G = e_1 \wedge (e_1 \leftrightarrow e_2 \vee p) \wedge (e_2 \leftrightarrow q \wedge r)$
- $F \equiv G$?
  No, because $I = \{p = 1, q = r = e_1 = e_2 = 0\}$ is a model of F and it is not a model of G

Exercises:

- Say wether the following statements are true or not:
  - For every formula F there is a logically equivalent formula G in CNF
  - For every formula F there is a unique logically equivalent formula G in CNF
  - If F is a tautology, there is a unique equivalent formula in CNF.
  - If F is a contradiction, there is a unique equivalent formula in CNF.

Exercises:

- Say wether the following statements are true or not:
  - For every formula F there is a logically equivalent formula G in CNF
    Yes. We just showed three ways to construct it.
  - For every formula F there is a unique logically equivalent formula G in CNF
    No. i.e, $(p \vee q) \wedge (p \vee \neg q) \equiv p$

Exercises:

- Say wether the following statements are true or not:
  - If F is a tautology, there is a unique equivalent formula in CNF.
    Yes. Only the empty formula (zero clauses) is a tautology.
    Proof: recall that in our definition of CNF we do not allow tautological clauses. Note that a non-tautological clause forbids at least one interpretation. Thus, if a CNF formula is not empty, it hast at least one interpretation which is not a model.
  - If F is a contradiction, there is a unique equivalent formula in CNF.
    No. i.e, $p \wedge \neg p$ and $p \vee q \wedge \neg p lor q \wedge p \vee \neg q \wedge \neg p lor \neg q$ are contradictions.

Final Remark

- Recall that our goal is to have a black-box SAT solver
- Now, we know that we can assume, without loss of generality, the the input is a CNF formula
- There are two foundamental approaches for SAT solvers:
  - Search
  - Inference
- The study of these two approaches will be the topic of the following sessions