Theories

We define a *theory* as a set of first-order sentences *closed under logical implication*.

Thus, $T$ is a theory iff $T$ is a set of sentences and if $T \models \sigma$, then $\sigma \in T$ for every sentence $\sigma$.

**Examples**

- For a given signature, the smallest possible theory consists of exactly the valid sentences over that signature.
- The largest theory for a given signature is the set of all sentences. It is the only unsatisfiable theory. Why?
Theories

For a class $\mathcal{K}$ of models over a given signature $\Sigma$, define the theory of $\mathcal{K}$ as

$$Th\mathcal{K} = \{ \sigma \mid \sigma \text{ is a } \Sigma\text{-sentence which is true in every model in } \mathcal{K} \}.$$ 

Theorem

$Th\mathcal{K}$ is indeed a theory.

Proof

Suppose $Th\mathcal{K} \models \sigma$. We know that $\models_M Th\mathcal{K}$ for each $M$ in $\mathcal{K}$. It follows that $\models_M \sigma$ for each $M$ in $\mathcal{K}$, and thus $\sigma \in Th\mathcal{K}$. 

Suppose $\Gamma$ is a set of sentences.

Define the set $Cn\ \Gamma$ of consequences of $\Gamma$ to be $\{ \sigma \mid \Gamma \models \sigma \}$. Then $Cn\ \Gamma = Th\ Mod\ \Gamma$. 
**Theories**

A theory $T$ is **complete** iff for every sentence $\sigma$, either $\sigma \in T$ or $(\neg \sigma) \in T$.

Note that if $M$ is a model, then $Th\{M\}$ is complete. In fact, for a class $\mathcal{K}$ of models, $Th\mathcal{K}$ is complete iff any two members of $\mathcal{K}$ are elementarily equivalent.

A theory $T$ is **axiomatizable** iff there is a decidable set $\Gamma$ of sentences such that $T = Cn\Gamma$.

A theory $T$ is **finitely axiomatizable** iff $T = Cn\Gamma$ for some finite set $\Gamma$ of sentences.

**Theorem**

If $Cn\Gamma$ is finitely axiomatizable, then there is a finite $\Gamma_0 \subseteq \Gamma$ such that $Cn\Gamma_0 = Cn\Gamma$.

**Proof**

If $Cn\Gamma$ is finitely axiomatizable, then for some sentence $\tau$, $Cn\Gamma = Cn\tau$.

Clearly, $\Gamma \models \tau$. By compactness, we have that there exists $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models \tau$. Thus, $Cn\tau \subseteq Cn\Gamma_0 \subseteq Cn\Gamma$, and since $Cn\Gamma = Cn\tau$, it follows that $Cn\Gamma_0 = Cn\Gamma$. 

\[ \square \]
Theories

Using the above terminology, we can restate our earlier results as follows:

- An axiomatizable theory (in a reasonable language) is effectively enumerable.
- A complete axiomatizable theory (in a reasonable language) is decidable.

Our results about theories can be summarized in the following diagram.
Los-Vaught Test

For a theory $T$ and a cardinal $\lambda$, say that $T$ is $\lambda$-categorical iff all models of $T$ having cardinality $\lambda$ are isomorphic.

**Theorem**

Let $T$ be a theory in a countable language such that

- $T$ is $\lambda$-categorical for some infinite cardinal $\lambda$.
- All models of $T$ are infinite.

Then $T$ is complete.

**Proof**

It suffices to show that for any two models $M$ and $M'$ of $T$, $M \equiv M'$. Since $M$ and $M'$ are infinite, there exist (by LST) elementarily equivalent models of cardinality $\lambda$. But these models must be isomorphic, and by the homomorphism theorem, isomorphic models are elementarily equivalent.
Validity and Satisfiability Modulo Theories

Given a $\Sigma$-theory $T$, a $\Sigma$-formula $\phi$ is

1. $T$-valid if $\models_M \phi[s]$ for all models $M$ of $T$ and all variable assignments $s$.
2. $T$-satisfiable if there exists some model $M$ of $T$ and variable assignment $s$ such that $\models_M \phi[s]$.
3. $T$-unsatisfiable if $\not\models_M \phi[s]$ for all models $M$ of $T$ and all variable assignments $s$.

The validity problem for $T$ is the problem of deciding, for each $\Sigma$-formula $\phi$, whether $\phi$ is $T$-valid.

The satisfiability problem for $T$ is the problem of deciding, for each $\Sigma$-formula $\phi$, whether $\phi$ is $T$-satisfiable.

Similarly, one can define the quantifier-free validity problem and the quantifier-free satisfiability problem for a $\Sigma$-theory $T$ by restricting the formula $\phi$ to be quantifier-free.
Validity and Satisfiability Modulo Theories

A decision problem is *decidable* if there exists an effective procedure which always terminates with an answer for any given instance of the problem.

For example, the validity problem for a $\Sigma$-theory $T$ is decidable if there exists an effective procedure for determining whether $T \models \phi$ for every $\Sigma$-formula $\phi$.

Note that validity problems can always be reduced to satisfiability problems:

$$\phi \text{ is } T\text{-valid iff } \neg\phi \text{ is } T\text{-unsatisfiable}.$$  

We will consider a few examples of theories which are of particular interest in verification applications.
The Theory $T_E$ of Equality

The theory $T_E$ of equality is the theory $Cn \emptyset$.

Note that the exact set of sentences in $T_E$ depends on the signature in question.

The theory does not restrict the possible values of symbols in any way. For this reason, it is sometimes called the theory of equality with uninterpreted functions (EUF).

The satisfiability problem for $T_E$ is just the satisfiability problem for first order logic, which is undecidable.

The satisfiability problem for conjunctions of literals in $T_E$ is decidable in polynomial time using congruence closure.
The Theory $T_{\mathbb{Z}}$ of Integers

Let $\Sigma_{\mathbb{Z}}$ be the signature $(0, 1, +, -, \leq)$.

Let $\mathcal{A}_{\mathbb{Z}}$ be the standard model of the integers with domain $\mathbb{Z}$.

Then $T_{\mathbb{Z}}$ is defined to be $Th\mathcal{A}_{\mathbb{Z}}$.

As showed by Presburger in 1929, the validity problem for $T_{\mathbb{Z}}$ is decidable, but its complexity is triply-exponential.

The quantifier-free satisfiability problem for $T_{\mathbb{Z}}$ is “only” NP-complete.

Let $\Sigma_{\mathbb{Z}}^\times$ be the same as $\Sigma_{\mathbb{Z}}$ with the addition of the symbol $\times$ for multiplication, and define $\mathcal{A}_{\mathbb{Z}}^\times$ and $T_{\mathbb{Z}}^\times$ in the obvious way.

The satisfiability problem for $T_{\mathbb{Z}}^\times$ is undecidable (a consequence of Gödel’s incompleteness theorem).

In fact, even the quantifier-free satisfiability problem for $T_{\mathbb{Z}}^\times$ is undecidable.
The Theory $T_R$ of Reals

Let $\Sigma_R$ be the signature $(0, 1, +, -, \leq)$.

Let $A_R$ be the standard model of the reals with domain $R$.

Then $T_R$ is defined to be $Th A_R$.

The satisfiability problem for $T_R$ is decidable, but the complexity is doubly-exponential.

The quantifier-free satisfiability problem for conjunctions of literals (atomic formulas or their negations) in $T_R$ is solvable in polynomial time, though exponential methods (like Simplex or Fourier-Motzkin) often perform better in practice.

Let $\Sigma^\times_R$ be the same as $\Sigma_R$ with the addition of the symbol $\times$ for multiplication, and define $A^\times_R$ and $T^\times_R$ in the obvious way.

In contrast to the theory of integers, the satisfiability problem for $T^\times_R$ is decidable.
The Theory $T_A$ of Arrays

Let $\Sigma_A$ be the signature $(\text{read}, \text{write})$.

Let $\Lambda_A$ be the following axioms:

1. $\forall a \forall i \forall v \left( \text{read} \left( \text{write} \left( a, i, v \right), i \right) = v \right)$
2. $\forall a \forall i \forall j \forall v \left( i \neq j \rightarrow \text{read} \left( \text{write} \left( a, i, v \right), j \right) = \text{read} \left( a, j \right) \right)$
3. $\forall a \forall b \left( (\forall i \left( \text{read} \left( a, i \right) = \text{read} \left( b, i \right) \right)) \rightarrow a = b \right)$

Then $T_A = Cn \Lambda_A$.

The satisfiability problem for $T_A$ is undecidable, but the quantifier-free satisfiability problem for $T_A$ is decidable (the problem is NP-complete).
Theories of Inductive Data Types

An *inductive data type* (IDT) defines one or more *constructors*, and possibly also *selectors* and *testers*.

**Example:** *list of int*

- **Constructors:** `cons : (int, list) → list, null : list`
- **Selectors:** `car : list → int, cdr : list → list`
- **Testers:** `is_cons, is_null`

The *first order theory* of a inductive data type associates a function symbol with each constructor and selector and a predicate symbol with each tester.

**Example:** `∀ x : list. (x = null ∨ ∃ y : int, z : list. x = cons(y, z))`

For IDTs with a single constructor, a conjunction of literals is decidable in polynomial time.

For more general IDTs, the problem is NP-complete, but reasonably efficient algorithms exist in practice.
Other Interesting Theories

Some other interesting theories include:

- Theory of bit-vectors
- Fragments of set theory
- Theory of floating-point arithmetic
- Theory of strings

SMT-LIB standard supports many different theories:
http://smtlib.cs.uiowa.edu/logics.shtml