syn.1 Formation Sequences

fol:syn:fseq:sec Defining formulas via an inductive definition, and the complementary technique of proving properties of formulas via induction, is an elegant and efficient approach. However, it can also be useful to consider a more bottom-up, step-by-step approach to the construction of formulas, which we do here using the notion of a *formation sequence*. To show how terms and formulas can be introduced in this way without needing to refer to their inductive definitions, we first introduce the notion of an arbitrary string of symbols drawn from some language $\mathcal{L}$.

fol:syn:fseq:defn:string **Definition syn.1 (Strings).** Suppose $\mathcal{L}$ is a first-order language. An $\mathcal{L}$-*string* is a finite sequence of symbols of $\mathcal{L}$. Where the language $\mathcal{L}$ is clearly fixed by the context, we will often refer to a $\mathcal{L}$-string simply as a *string*.

Example syn.2. For any first-order language $\mathcal{L}$, all $\mathcal{L}$-formulas are $\mathcal{L}$-strings, but not conversely. For example,

$$)(v_0 \to \exists$$

is an $\mathcal{L}$-string but not an $\mathcal{L}$-formula.

fol:syn:fseq:defn:fseq-trm **Definition syn.3 (Formation sequences for terms).** A finite sequence of $\mathcal{L}$-strings $\langle t_0, \dots, t_n \rangle$ is a *formation sequence* for a term t if $t \equiv t_n$ and for all $i \leq n$, either t_i is a variable or a constant symbol, or $\mathcal{L}$ contains a k-ary function symbol f and there exist $m_0, \dots, m_k < i$ such that $t_i \equiv f(t_{m_0}, \dots, t_{m_k})$.

Example syn.4. The sequence

$$\langle c_0, v_0, f_0^2(c_0, v_0), f_0^1(f_0^2(c_0, v_0)) \rangle$$

is a formation sequence for the term $f_0^1(f_0^2(c_0, v_0))$, as is

$$\langle v_0, c_0, f_0^2(c_0, v_0), f_0^1(f_0^2(c_0, v_0)) \rangle.$$

fol:syn:fseq:defn:fseq-frm **Definition syn.5 (Formation sequences for formulas).** A finite sequence of $\mathcal{L}$-strings $\langle \varphi_0, \dots, \varphi_n \rangle$ is a *formation sequence* for φ if $\varphi \equiv \varphi_n$ and for all $i \leq n$, either φ_i is an atomic formula or there exist $j, k < i$ and a variable x such that one of the following holds:

1. $\varphi_i \equiv \lnot\varphi_j$.
2. $\varphi_i \equiv (\varphi_j \land \varphi_k)$.
3. $\varphi_i \equiv (\varphi_j \lor \varphi_k)$.
4. $\varphi_i \equiv (\varphi_j \to \varphi_k)$.
5. $\varphi_i \equiv (\varphi_j \leftrightarrow \varphi_k)$.

6. $\varphi_i \equiv \forall x\, \varphi_j$.
7. $\varphi_i \equiv \exists x\, \varphi_j$.

Example syn.6.

$$\langle A^1_0(v_0), A^1_1(c_1), (A^1_1(c_1) \land A^1_0(v_0)), \exists v_0\, (A^1_1(c_1) \land A^1_0(v_0))\rangle$$

is a formation sequence of $\exists v_0\, (A^1_1(c_1) \land A^1_0(v_0))$, as is

$$\langle A^1_0(v_0), A^1_1(c_1), (A^1_1(c_1) \land A^1_0(v_0)), A^1_1(c_1),$$
$$\forall v_1\, A^1_0(v_0), \exists v_0\, (A^1_1(c_1) \land A^1_0(v_0))\rangle.$$

As can be seen from the second example, formation sequences may contain "junk": formulas which are redundant or do not contribute to the construction.

Proposition syn.7. *Every formula φ in $\mathrm{Frm}(\mathcal{L})$ has a formation sequence.* fol:syn:fseq:prop:formed

Proof. Suppose φ is atomic. Then the sequence $\langle\varphi\rangle$ is a formation sequence for φ. Now suppose that ψ and χ have formation sequences $\langle\psi_0, \dots, \psi_n\rangle$ and $\langle\chi_0, \dots, \chi_m\rangle$ respectively.

1. If $\varphi \equiv \lnot\psi$, then $\langle\psi_0, \dots, \psi_n, \lnot\psi_n\rangle$ is a formation sequence for φ.
2. If $\varphi \equiv (\psi \land \chi)$, then $\langle\psi_0, \dots, \psi_n, \chi_0, \dots, \chi_m, (\psi_n \land \chi_m)\rangle$ is a formation sequence for φ.
3. If $\varphi \equiv (\psi \lor \chi)$, then $\langle\psi_0, \dots, \psi_n, \chi_0, \dots, \chi_m, (\psi_n \lor \chi_m)\rangle$ is a formation sequence for φ.
4. If $\varphi \equiv (\psi \to \chi)$, then $\langle\psi_0, \dots, \psi_n, \chi_0, \dots, \chi_m, (\psi_n \to \chi_m)\rangle$ is a formation sequence for φ.
5. If $\varphi \equiv (\psi \leftrightarrow \chi)$, then $\langle\psi_0, \dots, \psi_n, \chi_0, \dots, \chi_m, (\psi_n \leftrightarrow \chi_m)\rangle$ is a formation sequence for φ.
6. If $\varphi \equiv \forall x\, \psi$, then $\langle\psi_0, \dots, \psi_n, \forall x\, \psi_n\rangle$ is a formation sequence for φ.
7. If $\varphi \equiv \exists x\, \psi$, then $\langle\psi_0, \dots, \psi_n, \exists x\, \psi_n\rangle$ is a formation sequence for φ.

By the principle of induction on formulas, every formula has a formation sequence. □

We can also prove the converse. This is important because it shows that our two ways of defining formulas are equivalent: they give the same results. It also means that we can prove theorems about formulas by using ordinary induction on the length of formation sequences.

Lemma syn.8. *Suppose that $\langle\varphi_0, \dots, \varphi_n\rangle$ is a formation sequence for φ_n, and that $k \le n$. Then $\langle\varphi_0, \dots, \varphi_k\rangle$ is a formation sequence for φ_k.* fol:syn:fseq:lem:fseq-init

Proof. Exercise. □

Problem syn.1. Prove Lemma syn.8.

fol:syn:fseq: thm:fseq-frm-equiv

Theorem syn.9. $\mathrm{Frm}(\mathcal{L})$ *is the set of all expressions (strings of symbols) in the language* $\mathcal{L}$ *with a formation sequence.*

Proof. Let F be the set of all strings of symbols in the language $\mathcal{L}$ that have a formation sequence. We have seen in Proposition syn.7 that $\mathrm{Frm}(\mathcal{L}) \subseteq F$, so now we prove the converse.

Suppose φ has a formation sequence $\langle \varphi_0, \dots, \varphi_n \rangle$. We prove that $\varphi \in \mathrm{Frm}(\mathcal{L})$ by strong induction on n. Our induction hypothesis is that every string of symbols with a formation sequence of length $m < n$ is in $\mathrm{Frm}(\mathcal{L})$. By the definition of a formation sequence, either $\varphi \equiv \varphi_n$ is atomic or there must exist $j, k < n$ such that one of the following is the case:

1. $\varphi \equiv \neg\varphi_j$.
2. $\varphi \equiv (\varphi_j \land \varphi_k)$.
3. $\varphi \equiv (\varphi_j \lor \varphi_k)$.
4. $\varphi \equiv (\varphi_j \rightarrow \varphi_k)$.
5. $\varphi \equiv (\varphi_j \leftrightarrow \varphi_k)$.
6. $\varphi \equiv \forall x\, \varphi_j$.
7. $\varphi \equiv \exists x\, \varphi_j$.

Now we reason by cases. If φ is atomic then $\varphi_n \in \mathrm{Frm}(\mathcal{L}_0)$. Suppose instead that $\varphi \equiv (\varphi_j \land \varphi_k)$. By Lemma syn.8, $\langle \varphi_0, \dots, \varphi_j \rangle$ and $\langle \varphi_0, \dots, \varphi_k \rangle$ are formation sequences for φ_j and φ_k, respectively. Since these are proper initial subsequences of the formation sequence for φ, they both have length less than n. Therefore by the induction hypothesis, φ_j and φ_k are in $\mathrm{Frm}(\mathcal{L}_0)$, and by the definition of a formula, so is $(\varphi_j \land \varphi_k)$. The other cases follow by parallel reasoning. □

Formation sequences for terms have similar properties to those for formulas.

fol:syn:fseq: prop:fseq-trm-equiv

Proposition syn.10. $\mathrm{Trm}(\mathcal{L})$ *is the set of all expressions* t *in the language* $\mathcal{L}$ *such that there exists a (term) formation sequence for* t.

Proof. Exercise. □

Problem syn.2. Prove Proposition syn.10. Hint: use a similar strategy to that used in the proof of Theorem syn.9.

There are two types of "junk" that can appear in formation sequences: repeated elements, and elements that are irrelevant to the construction of the formation or term. We can eliminate both by looking at minimal formation sequences.

Definition syn.11 (Minimal formation sequences). A formation sequence $\langle \varphi_0, \dots, \varphi_n \rangle$ for φ is a *minimal formation sequence* for φ if for every other formation sequence s for φ, the length of s is greater than or equal to $n + 1$. fol:syn:fseq: defn:minimal-fseq

Proposition syn.12. *The following are equivalent:* *fol:syn:fseq: prop:subformula-equivs*

1. *ψ is a sub-formula of φ.*
2. *ψ occurs in every formation sequence of φ.*
3. *ψ occurs in a minimal formation sequence of φ.*

Proof. Exercise. □

Problem syn.3. Prove Proposition syn.12.

Historical Remarks Formation sequences were introduced by Raymond Smullyan in his textbook *First-Order Logic* (Smullyan, 1968). Additional properties of formation sequences were established by Zuckerman (1973).

Photo Credits

Bibliography

Smullyan, Raymond M. 1968. *First-Order Logic*. New York, NY: Springer. Corrected reprint, New York, NY: Dover, 1995.

Zuckerman, Martin M. 1973. Formation sequences for propositional formulas. *Notre Dame Journal of Formal Logic* 14(1): 134–138.