Logic of feature and feature-structure spaces

D. Terence Langendoen
Dept of Linguistics, University of Arizona & Linguistics Program, National Science Foundation
A Boolean feature (BF) space is ...

1. a finite set $F$ of features (attribute-value pairs, sharing a common attribute), and

2. a partial ordering relation $\sqsubseteq$ on the values of $F$ (subsumption).
   - $F$ contains a subset of at least two atoms (generators) that do not subsume each other (are logically independent).
   - $F$ is closed under the logical operators:
     & Conjunction (same as “unification”)  
     | Disjunction (same as “generalization”)  
     ~ Negation (same as “opposition”)
Subsumption is ...

- Reflexive
  - $f \sqsupseteq f$ for all $f \in F$

- Antisymmetric
  - if $f \sqsubseteq g$ and $g \sqsubseteq f$, then $f = g$

- Transitive
  - if $f \sqsubseteq g$ and $g \sqsubseteq h$, then $f \sqsubseteq h$

- Monotonic
  - if $f \sqsubseteq g$, then $f, h \sqsubseteq g$ for any $h \in F$
A BF space with two generators

• Let $G_2$ be the BF space with the generators \{f, m\}, where:
  \( f = [\text{Gender feminine}] \)
  \( m = [\text{Gender masculine}] \)
• Then $G_2$ also contains:
  \( f \mid m = [\text{Gender (feminine | masculine)}] = \top \) (top, or “verum”), and
  \( f \& m = [\text{Gender (feminine & masculine)}] = \bot \) (bottom, or “falsum”)
• Moreover:
  \( \sim f = [\text{Gender \sim feminine}] = m \quad \sim m = f \quad \sim \top = \bot \quad \sim \bot = \top \)
• The next slide shows the subsumption relation among the four $G_2$ features, omitting reflexive arcs and arcs derivable from transitivity.
Note: Read the arcs downwards for conjunction (unification), and upwards for disjunction (generalization). In BF spaces, the negation (opposition) of a feature is the feature appearing in its “reflection” using the x and y axes as mirrors.
A BF space with three generators

• Let $G_3$ be the BF space with the three generators $\{f, m, n\}$, where $f$ and $m$ are as in $G_2$ and:
  \[ n = \text{[Gender neuter]} \]

• The remaining five features in $G_3$ are:
  \[
  f \mid m = \neg n \quad f \mid n = \neg m \quad m \mid n = \neg f \quad \top \quad \bot
  \]

• The next slide shows the subsumption relation among the eight $G_3$ features.
Size of a BF space depends on the number of generators for that space

- Let $n$ be the number of generators in a BF space $F$.
- Then the number of features in $F$ is $2^n$.
- The number of features in row $i$ (counting in either direction) in a subsumption diagram of $F$ is given by: $n! / [(i - 1)! * (n - i + 1)!]$
- So the BF number spaces $N_2$ and $N_3$ with generators $\{sg, pl\}$ and $\{sg, du, mu\}$ look just like the BF spaces $G_2$ and $G_3$ with respect to subsumption.
Note: sg = [Number singular], pl = [Number plural]
Note: du = [Number dual], mu = [Number multal]
Reduced feature (RF) spaces

• Some features in a BF space may not be realized in a language because they are not associated with any expression in that language, either:
  – directly by lexical specification, or
  – as a result of the application of a grammatical rule.
• Removing such features (except for ⊥, retained to express unification failure) from a BF space results in a reduced feature (RF) space.
• I leave open the question of why certain features are realized and others not for another occasion.
Three examples of RF spaces

1. Language $L_1$ realizes the $G_3$ features:
   \[ f, m, n, m|f = c \text{ ("common")}, \text{ and } \top; \text{ but not } m|n, f|n. \]

2. Language $L_2$ realizes the $N_3$ features:
   \[ \text{sg, du, mu, sg|du = pa ("paucal"), and du|mu = pl, but not sg|mu or } \top. \]

3. Language $L_3$ is like $L_2$, except:
   it realizes $\top$ and does not realize mu.

- The resulting RF spaces ($G_{3a}, N_{3a}, N_{3b}$) are shown in the next three slides.
$N_{3a}$

Diagram:

```
\[
\begin{array}{ccc}
pa & & pl \\
\downarrow & & \uparrow \\
sg & & du & & mu \\
\downarrow & & \uparrow \\
& & \bot & & \\
\end{array}
\]
Non-Boolean properties of $G_{3a}$ and $N_{3a}$

- Negation is not everywhere defined in $G_{3a}$ and $N_{3a}$.
  - In $G_{3a}$, $\neg n = c$ and conversely, but both $\neg f$ and $\neg m$ are undefined.
    - Both $m$ and $c$ are candidates for $\neg f$, but there is no basis for choosing between them; similarly for $\neg m$.
  - $\neg sg = pl$ and $\neg mu = pa$ and conversely, but $\neg du$ and $\bot$ are undefined.
    - Both $sg$ and $mu$ are candidates for $\neg du$, but there is no basis for choosing between them; similarly for $pa$ and $pl$ as candidates for $\neg \bot$.
- Disjunction is not everywhere defined in $G_{3b}$.
  - In $G_{3b}$, $sg|du = pa$ and $du|mu = pl$, but $sg|mu$ is undefined.
    - Both $pa$ and $pl$ are candidates for $sg|mu$, but there is no basis for choosing between them.
Non-Boolean properties of $N_{3b}$

- In $N_{3b}$, negation and disjunction are everywhere defined, but the laws of double negation (NN) and excluded middle (XM) that hold in BF spaces fail.
  - While $\sim\sim\text{sg} = \sim\text{pl} = \text{sg}$ (satisfying NN), $\sim\text{du} = \text{sg}$, so that $\sim\sim\text{du} = \sim\text{sg} = \text{pl}$, not $\text{du}$ (thus failing NN).
  - Similarly, while $\text{sg} \mid \sim\text{sg} = \text{sg} \mid \text{pl} = \top$ (satisfying XM), $(\text{sg}\mid\text{du}) \mid \sim(\text{sg}\mid\text{du}) = (\text{sg}\mid\text{du}) \mid \bot = \text{sg}\mid\text{du}$, not $\top$ (thus failing XM).
A Boolean feature-structure (BFS) space is ...

- the Cartesian product of two or more BF spaces.
  - For example, the BFS space $G_2 \times N_2$ has the four generators \{f$\times$sg, m$\times$sg, f$\times$pl, m$\times$pl\}, hence 16 members, related by subsumption as in the next slide.
  - Note that the FSs sg = $\top$×sg, f = f$\times$⊤, ⊤ = ⊤×⊤, etc.
- Explicit disjunctions of feature structures are shown in italics.
  - Note that f $\leftrightarrow$ sg = (f$\times$sg) $|$ (m$\times$pl), similarly for m $\leftrightarrow$ sg.
$G_2 \times N_2$

Diagram showing the relationships between various elements labeled as $f|sg$, $m|sg$, $f|pl$, $m|pl$, $sg$, $f\leftrightarrow sg$, $f$, $m$, $m\leftrightarrow sg$, $pl$, $f\times sg$, $m\times sg$, $f\times pl$, $m\times pl$. The diagram illustrates the logical connections and interactions between these elements within the context of feature and feature-structure spaces.
Reduced feature-structure (RFS) spaces

- Even if all the features (other than ⊥) in some BFS space are realized in a language, some feature structures in that space may not be realized in that language.
  - As with features, I am not concerned here with the problem of explaining why certain FSs are realized and not others.
  - Removing unrealized FSs (again, except for ⊥) from a BFS space results in a reduced feature-structure (RFS) space.
Example of an RFS space

• For example, consider a grammar that lexicalizes some of the generators of $G_2 \times N_2$, along with the underspecified FSs $\top, f, m, sg$, and $pl$, and that realizes all of the remaining generators via unification.
• Then the RFS space $(G_2 \times N_2)_a$ associated with that grammar is as in the next slide.
• The disjunctive FSs in the original BFS space (indicated in italics in the diagram of $G_2 \times N_2$) have been removed.
\[(G_2 \times N_2)_a\]
Non-Boolean character of some RFS spaces

• An RFS space resulting from reducing a BFS space is generally non-Boolean.
  – In particular, the negations of the generators in the \((G_2 \times N_2)_a\) space are not defined.
  • For example, there are two candidates for \(\sim(f \times sg)\), namely m and pl, and there is no basis for choosing one over the other.
RFS spaces derived from non-Boolean FS spaces

- A FS space which is the Cartesian product of RF spaces may be further reduced.
  - For example, let $N_{3c}$ be the same as $N_{3b}$, but with the removal of $pa$, and consider the FS space $G_{2a} \times N_{3c}$.
  - Let $L_4$ be the language realizing all the FSs in that space, except for the generator $n \times du$, the explicitly disjunctive FSs and the FS du.
  - The resulting non-Boolean RFS space $(G_{2a} \times N_{3c})_a$ contains 19 FSs with the subsumption relations shown in the next slide.
\((G_{2a} \times N_{3c})_a\)