Proportional analogies in typed systems (HPSG?)

Matías Guzmán Naranjo - Universität Düsseldorf

EW-HPSG
02.06.2018
Proportional analogies
What are PA models

In PA (proportional analogy) models of morphology, the different inflected forms of a word are not produced by inflectional rules, but simply associated by analogical relations.

A proportional analogy is expressed by the formula:

\[ A : B :: C : X \]  

meaning A is to B like C is to X. Where if we know A, B and C, we can deduce X.
What are PA models

As an example:

(1)   a. compr**o**
      buy.1sg

       b. compras
       buy.2sg

(2)   a. paro
      stop.1sg

       b. X
       stop.2sg

**compro:**compras::paro:**X**

⇒ X=**paras**
What are PA models

But also:

\[
\begin{array}{ccc}
\text{amár} & : & \text{amará} & : & \text{amára} \\
\text{parár} & : & \text{parará} & : & \text{parára} \\
\text{LOVE.INF} & : & \text{LOVE.3SG.FUT.IND} & : & \text{LOVE.1/3SG.PRES.SUBJ} \\
\end{array}
\]

For which there is no obvious morphemic segmentation, or function/semantic relation.
What are PA models

We can express analogies between paradigm cells by abstracting away the phonological material common to those cells:

\[
\begin{array}{ccc}
X' & : & Xá' \\
\text{INF} & : & \text{3SG.FUT.IND} \\
\end{array}
\]

These are not ‘rules’ because the \( X \) and the segment \( a \) does not correspond to a morpheme and there is no feature composition between these cells.

I will call this the ‘X-notation’
Advantages of PA models

There are several advantages of PA models:

- no need for morphemes
- (thus) no segmentation problem
- no transformations,
- no special rules,
- no rule orderings,
- no feature decompositions.

In short: they are very simple and intuitive.
Proportional analogy

Problems with PA models

And then there are problems (at least with the X-notation):

- there is no clear interface between the morphology and the rest of the grammar,
- only one level of abstraction (e.g. no abstraction across inflection classes),
- their status is unclear (what exactly are proportional analogies and how do they work⁉),
- not properly formalized,
- hard(er) to implement as a complete system,
- incompatible with theories that require lexemes (or not? hard to tell).
Proportional analogy

Problems with PA models

A concrete example with some Spanish verb cells:

<table>
<thead>
<tr>
<th></th>
<th>-ar</th>
<th>-er</th>
<th>-ir</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>regular</td>
<td>o-ue</td>
<td>regular</td>
</tr>
<tr>
<td>infinitive</td>
<td>cant-ár</td>
<td>prob-ár</td>
<td>com-ér</td>
</tr>
<tr>
<td>1sg present</td>
<td>cánt-o</td>
<td>pruéb-o</td>
<td>cóm-o</td>
</tr>
<tr>
<td>2sg (tú) present</td>
<td>cánt-as</td>
<td>pruéb-as</td>
<td>cóm-es</td>
</tr>
<tr>
<td>2sg (vos) present</td>
<td>cant-ás</td>
<td>prob-ás</td>
<td>com-és</td>
</tr>
<tr>
<td>3sg future</td>
<td>cant-ará</td>
<td>prob-ará</td>
<td>com-erá</td>
</tr>
<tr>
<td>1sg imp.subj</td>
<td>cant-ára</td>
<td>prob-ára</td>
<td>com-iéra</td>
</tr>
</tbody>
</table>

|                 | ‘sing’ | ‘taste’ | ‘eat’ | ‘hurt’ | ‘go up’ | ‘die’ |

We can analyze the Spanish verbal inflection system as being composed of a stem (process), a stress pattern and a suffix or set of suffixes.
Proportional analogy

Problems with PA models

We can define concrete analogical relations between any two cells in the previous examples.

(5)  
   a. amár:ámas::cantár:X
   b. Yr:Ys
   c. X = cántas

However, we cannot properly specify analogies covering more general patterns. For example, the analogy in (5) only covers *ar* and *er* verbs, but not *ir* verbs: *subir*-subes

However, we miss the fact that for all three classes the second person singular present indicative has an -s marker.
Proportional analogy

Problems with PA models

The difficulty with analogies is that it is not clear how to express independent partial patterns (stress, stem alternations, suffix markers), and then combine these individual patterns together. A naive approach would not work. The stem alternations alone represent a problem:

(6)  a. Xar:Xo
    b. XoYar:XueYo

(6a) cannot capture stem alternating verbs, while (6b) cannot capture non-alternating verbs. Similarly, a partial analogy only for stem alternations as in (7) could not be unified with a partial analogy like in (6a).

(7)  a. XoY:XueY

Moreover, (7) is exclusively making reference to stems (it does not mention suffix markers), but PA approaches do not assume stems.
Proportional analogy

Problems with PA models

Some generalizations we would like to express, but cannot:

- -o marks first person singular present indicative,
- stress patterns are identical across inflection classes,
- stem-Vs marks second person singular,
- some verbs share exactly the same stem alternation process even though they belong to different thematic vowels.
Proportional analogy

Problems with PA models

Without proper formalization it is not clear how we can solve these issues, and whether these are limitations of the notation or of PAs themselves.
Objectives
Objectives

This talk has aims to present a path to the formalization of a PA model which:

- retains he simplicity characteristic of analogical models (no rules, orderings, stems, morphemes, etc.),
- is completely surface oriented,
- allows for partial analogies,
- allows for lexemes,
- could potentially interact with syntactic theories,
- and lets us test the limits of analogy.
Basic assumptions
Basic assumptions

The basic assumptions are:

- Lexemes bundle together all their possible realizations.
- Analogical relations hold across phonological strings, not words.
- Analogical relations are constrains on phonological strings.
- Analogical relations are organized in an inheritance hierarchy according to degree of abstraction.
The basics

Where FTS contains whatever features are associated with any given cell in the paradigm (syntax, agreement, semantics, pragmatics, etc.), and PHON a phonological string (whatever its representation may be).
Phonological representation?

\[
\begin{align*}
\text{phon-obj} & \quad \text{WORD list(syllable-obj)} \\
\text{syllable} & \quad \text{ONSET} \ldots \\
\text{RIME} & \quad \text{NUCLEUS} \ldots \\
\text{CODA} & \quad \text{vowel} \\
& \quad \text{HIGH } +/- \\
& \quad \text{ROUND } +/- \\
& \quad \text{BACK } +/- 
\end{align*}
\]
Some analogies

cantar ‘to sing’

\[
\text{lexeme-cantar} = \begin{cases} 
\text{FTS 1sg.pres.ind PHON }/\#\text{canto}#/\text{,} \\
\text{FTS 2sg.pres.ind PHON }/\#\text{cantas}#/\text{,} \\
\text{FTS 3sg.pres.ind PHON }/\#\text{canta}#/\text{,} \\
\text{FTS 1pl.pres.ind PHON }/\#\text{cantamos}#/\text{,} \\
\text{FTS 2pl.pres.ind PHON }/\#\text{cantáis}#/\text{,} \\
\text{FTS 3pl.pres.ind PHON }/\#\text{cantan}#/\text{,} \\
\end{cases}
\]

And similarly for comer and subir.
Ideally, however, we would like to define objects between the fully abstract *lexeme* type, and the fully specified *lexeme-cantar* type.

The idea of PA models is that we can reconstruct missing forms in a paradigm.

We do this by abstracting away the proportional analogies across all phonological strings in the paradigm of *cantar, comer, subir*, etc.
Some analogies

## Analogies

<table>
<thead>
<tr>
<th>ar-class</th>
<th>er-class</th>
<th>ir-class</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="ar-class-tree.png" alt="Tree" /></td>
<td><img src="er-class-tree.png" alt="Tree" /></td>
<td><img src="ir-class-tree.png" alt="Tree" /></td>
</tr>
</tbody>
</table>

*FTS* 1sg.prs  PH / $X^1$ o#/

*FTS* 2sg.prs  PH / $X^1$ as#/

*FTS* 3sg.prs  PH / $X^1$ a#/  

*FTS* 1pl.prs  PH / $X^2$ amos#/  

*FTS* 2pl.prs  PH / $X^2$ ais#/

*FTS* 3pl.prs  PH / $X^1$ an#/  

*FTS* 1sg.prs  PH / $X^1$ o#/  

*FTS* 2sg.prs  PH / $X^1$ es#/  

*FTS* 3sg.prs  PH / $X^1$ e#/  

*FTS* 1pl.prs  PH / $X^2$ emos#/  

*FTS* 2pl.prs  PH / $X^2$ eis#/

*FTS* 3pl.prs  PH / $X^1$ en#/  

*FTS* 1pl.prs  PH / $X^2$ imos#/  

*FTS* 2pl.prs  PH / $X^2$ is#/

*FTS* 3pl.prs  PH / $X^1$ en#/  

...
Some analogies

Analogies
Some analogies

Analogies

\[
\begin{array}{c}
\text{stress} \\
\begin{cases}
\text{FTS }1\text{sg-pres PHON }/\dot{o}\sigma#/ \\
\text{FTS }2\text{sg-pres PHON }/\dot{o}\sigma#/ \\
\text{FTS }3\text{sg-pres PHON }/\dot{o}\sigma#/ \\
\text{FTS }1\text{pl-pres PHON }/\dot{o}\sigma#/ \\
\text{FTS }2\text{pl-pres PHON }/\dot{o}#/ \\
\text{FTS }3\text{pl-pres PHON }/\dot{o}\sigma#/ \\
\text{...}
\end{cases}
\end{array}
\]

But since this pattern applies to all (*) verbs, we can think of it as a supper type of the individual patters for markers.
Some analogies

Analogies

\[\begin{array}{ll}
\text{stem-o-ue} \\
\text{FTS 1sg-p PHON} & /X^1 \sigma_{ue}^2 \sigma_{O1}/ \\
\text{FTS 2sg-p PHON} & /X^1 \sigma_{ue}^2 \sigma_{O1}/ \\
\text{FTS 3sg-p PHON} & /X^1 \sigma_{ue}^2 \sigma_{O1}/ \\
\text{FTS 1pl-p PHON} & /X^1 \sigma_{O} \sigma_{O1} \sigma/ \\
\text{FTS 2pl-p PHON} & /X^1 \sigma_{O} \sigma_{O1} \sigma/ \\
\text{FTS 3pl-p PHON} & /X^1 \sigma_{ue}^2 \sigma_{O1}/ \\
\end{array}\]

\[\begin{array}{ll}
\text{stem-regular} \\
\text{FTS 1sg-p PHON} & /X^1 \sigma_{O1}/ \\
\text{FTS 2sg-p PHON} & /X^1 \sigma_{O1}/ \\
\text{FTS 3sg-p PHON} & /X^1 \sigma_{O1}/ \\
\text{FTS 1pl-p PHON} & /X^1 \sigma_{O1} \sigma/ \\
\text{FTS 2pl-p PHON} & /X^1 \sigma_{O1} \sigma/ \\
\text{FTS 3pl-p PHON} & /X^1 \sigma_{O1}/ \\
\end{array}\]
Finally, we need a hierarchy to put everything together.
But we can do more...

The issue is that these examples are easily captured with any theory. There are two aspects of PA systems, however, which are rarely if ever captures by constructive approaches. These are predictability/information and relations between cells.

The first has to do with the fact that knowing one cell of a paradigm might give us partial or complete information about the other cells in the paradigm. Similarly, knowing part of a cell in a paradigm might give us information about which cell it is.

The second issue has to do with the fact that some generalizations are better expressed as relations between two fully inflected cells, and not as a derivation starting from a stem.
Some analogies

Partial predicatibility

<table>
<thead>
<tr>
<th>PARADIGM</th>
<th>FTS</th>
<th>PHON</th>
</tr>
</thead>
<tbody>
<tr>
<td>o-marker</td>
<td>1.present</td>
<td>/σ_{NUC/o}/#</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1-marker</th>
<th>FTS</th>
<th>PHON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg.present</td>
<td>/o#/</td>
<td></td>
</tr>
<tr>
<td>1pl.present</td>
<td>/mos#/</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1SG.</th>
<th>amo</th>
<th>como</th>
<th>siento</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PL.</td>
<td>amamos</td>
<td>comemos</td>
<td>sentimos</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARADIGM</th>
<th>FTS</th>
<th>PHON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PL.</td>
<td>amo</td>
<td>como</td>
</tr>
<tr>
<td>1PL.</td>
<td>amamos</td>
<td>comemos</td>
</tr>
</tbody>
</table>
## Cell relations

<table>
<thead>
<tr>
<th></th>
<th>cantar</th>
<th>eat</th>
<th>go up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INF</strong></td>
<td>cantár</td>
<td>comer</td>
<td>subír</td>
</tr>
<tr>
<td><strong>1SG.COND</strong></td>
<td>cantaría</td>
<td>comería</td>
<td>subiría</td>
</tr>
<tr>
<td><strong>1SG.FUT</strong></td>
<td>cantaré</td>
<td>comeré</td>
<td>subiré</td>
</tr>
<tr>
<td><strong>1SG.IMP.SBJ</strong></td>
<td>cantára</td>
<td>comiéra</td>
<td>subiéra</td>
</tr>
<tr>
<td><strong>3SG.FUT</strong></td>
<td>cantarás</td>
<td>comerás</td>
<td>subirás</td>
</tr>
<tr>
<td><strong>3SG.IMP.SBJ</strong></td>
<td>cantáras</td>
<td>comiéras</td>
<td>subiéras</td>
</tr>
<tr>
<td><strong>3SG.PRES</strong></td>
<td>cánta</td>
<td>cóme</td>
<td>súbe</td>
</tr>
<tr>
<td><strong>2SG.IMP/tú</strong></td>
<td>cánta</td>
<td>cóme</td>
<td>súbe</td>
</tr>
<tr>
<td><strong>2SG.IMP/vos</strong></td>
<td>cantá</td>
<td>comé</td>
<td>subí</td>
</tr>
<tr>
<td><strong>2SG.IMP/usted</strong></td>
<td>cánte</td>
<td>cóma</td>
<td>súba</td>
</tr>
<tr>
<td><strong>1SG.PRET</strong></td>
<td>canté</td>
<td>comí</td>
<td>subí</td>
</tr>
<tr>
<td><strong>1SG.PRES</strong></td>
<td>cánto</td>
<td>cómo</td>
<td>súbo</td>
</tr>
<tr>
<td><strong>3SG.PRET</strong></td>
<td>cantó</td>
<td>comió</td>
<td>subió</td>
</tr>
</tbody>
</table>
Some analogies

## Cell relations

<table>
<thead>
<tr>
<th>1SG.IMP.SUBJ</th>
<th>2SG.IMP.SUBJ</th>
<th>3SG.IMP.SUBJ</th>
<th>1PL.IMP.SUBJ</th>
<th>2PL.IMP.SUBJ</th>
<th>3PL.IMP.SUBJ</th>
<th>1SGCOND</th>
<th>2SGCOND</th>
<th>3SGCOND</th>
<th>1PLCOND</th>
<th>2PLCOND</th>
<th>3PLCOND</th>
</tr>
</thead>
<tbody>
<tr>
<td>cantar</td>
<td>cantára</td>
<td>comiéra</td>
<td>subiéra</td>
<td>cantára</td>
<td>cantáras</td>
<td>comiéras</td>
<td>subiéras</td>
<td>cantára</td>
<td>comiéra</td>
<td>subiéra</td>
<td>cantára</td>
</tr>
</tbody>
</table>
Morphemes

In the imperfect:

<table>
<thead>
<tr>
<th></th>
<th>cantar</th>
<th>comer</th>
<th>subir</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG.IMP</td>
<td>cant-ába</td>
<td>com-ía</td>
<td>sub-ía</td>
</tr>
<tr>
<td>2.SG.IMP</td>
<td>cant-ába-s</td>
<td>com-ía-s</td>
<td>sub-ía-s</td>
</tr>
<tr>
<td>3.SG.IMP</td>
<td>cant-ába</td>
<td>com-ía</td>
<td>sub-ía</td>
</tr>
<tr>
<td>1.PL.IMP</td>
<td>cant-ába-mos</td>
<td>com-ía-mos</td>
<td>sub-ía-mos</td>
</tr>
<tr>
<td>2.PL.IMP</td>
<td>cant-ába-is</td>
<td>com-ía-is</td>
<td>sub-ía-is</td>
</tr>
<tr>
<td>3.PL.IMP</td>
<td>cant-ába-n</td>
<td>com-ía-n</td>
<td>sub-ía-n</td>
</tr>
</tbody>
</table>
Some analogies

Morphemes

The same pattern repeats for all but the indefinido:

<table>
<thead>
<tr>
<th></th>
<th>cantar</th>
<th>comer</th>
<th>subir</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.IND</td>
<td>cant-é</td>
<td>com-í</td>
<td>sub-íó</td>
</tr>
<tr>
<td>2SG.IND</td>
<td>cant-asté</td>
<td>com-isté</td>
<td>sub-isté</td>
</tr>
<tr>
<td>3SG.IND</td>
<td>cant-ó</td>
<td>com-ió</td>
<td>sub-íó</td>
</tr>
<tr>
<td>1PL.IND</td>
<td>cant-a-<strong>mos</strong></td>
<td>com-i-<strong>mos</strong></td>
<td>sub-i-<strong>mos</strong></td>
</tr>
<tr>
<td>2PL.IND</td>
<td>cant-asté-<strong>is</strong></td>
<td>com-isté-<strong>is</strong></td>
<td>sub-isté-<strong>is</strong></td>
</tr>
<tr>
<td>3PL.IND</td>
<td>cant-aró-<strong>n</strong></td>
<td>com-ieró-<strong>n</strong></td>
<td>sub-ieró-<strong>n</strong></td>
</tr>
</tbody>
</table>

Which nonetheless retains the same markers for the plural (*mos, is, n*).
Some analogies

Morphemes

tam-markers

\[
\begin{align*}
\text{FTS} & \quad 1pl \\
\text{PHON} & \quad /#X^1 \text{mos}/
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 2pl \\
\text{PHON} & \quad /#X^1 \text{is}/
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 3pl \\
\text{PHON} & \quad /#X^1 n/
\end{align*}
\]

subj-imp-conditional

\[
\begin{align*}
\text{FTS} & \quad 1sg.(\text{subj} \lor \text{ind.imp} \lor \text{cond}) \\
\text{PHON} & \quad /#X^1 /
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 2sg.(\text{subj} \lor \text{ind.imp} \lor \text{cond}) \\
\text{PHON} & \quad /#X^1 s/
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 3sg.(\text{subj} \lor \text{ind.imp} \lor \text{cond}) \\
\text{PHON} & \quad /#X^1 /
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad \text{pl.}(\text{subj} \lor \text{ind.imp} \lor \text{cond}) \\
\text{PHON} & \quad /#X^1 /
\end{align*}
\]

imperfective

\[
\begin{align*}
\text{FTS} & \quad 2sg.imperf \\
\text{PHON} & \quad /s/
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 3/1sg.imperf \\
\text{PHON} & \quad /ba/
\end{align*}
\]

\[
\begin{align*}
\text{FTS} & \quad 3/1sg.imperf \\
\text{PHON} & \quad /ia/
\end{align*}
\]

imp-ar

imp-er-ir
Some analogies

**Morphemes**

```
paradigm
  ├── tam-markers
  │    ├── perfective
  │    └── imperfective
  └── cel-relations
    └── subj-imp-cond
        └── imperfect
            ├── imp-ar
            └── imp-er-ir
```
We can express both ‘morphemes’ and cell relations with the same technique, at the same time, and at the same level of abstraction.
A few issues

Some difficult issues I can see so far:

- It is hard to model partial overabundance
- Morphemes with variable positions are tricky
Complete overabundance is easy

The imperfect subjunctive in Spanish has two possible realizations: -se and -ra. This applies to all verbs:

(8)  
   a. comiera - comiese ‘eat’  
   b. supiera - supiese ‘know’  
   c. amara - amase ‘love’  
   d. ...

We can model this by simply having several cells with identical morphosyntactic features but different phonological constraints:

{[FTS 1sg.imp.subj PHON /#Xra#/],  
 [FTS 1sg.imp.subj PHON /#Xse#/]}
Partial overabundance is harder

The issue arises with systems in which some items, but not all, are overabundant:

The Spanish plural system is relatively trivial:

(9)  
    a. gato - gatos (‘cat’)
    b. perro - perros (‘dog’)
    c. árbol - árboles (‘tree’)
    d. baúl - baules (‘chest’)

With some exceptions:

(10)  
    a. ají - ajís ~ ajíes (‘chili pepper’)
    b. colibrí - colibrís ~ colibríes (‘hummingbird’)

Overabundance as multiple inheritance

An intuitive way to think about overabundance is as multiple inheritance:

```
plural
\|-- s
   \|-- es
      \|-- pl-s
         \|-- lxm-gato
      \|-- pl-es\~s
          \|-- lxm-ají
      \|-- pl-es
          \|-- lxm-árbol
```

But with normal unification we cannot grow this paradigm.
What to do?

• A way around this is unification as in CxG (with set union), but...?
• maybe an exception just for paradigms...?
• having three independent plural classes seems **wrong** because it misses the fact one class is the combination of the other two.
Naming cells?

Something like:

```
PARADIGM
[ SINGULAR [ FTS sg PHON ... ] ]
[ PLURAL-1 [ FTS pl PHON ... ] ]
[ PLURAL-2 [ FTS pl PHON ... ] ]
```

Also seems quite wrong.
But maybe:

```
number
  /   
pl-s  pl-es
     /   /   
  lxm-s-gato  lxm-s-ají  lxm-es-ají  lxm-es-árbol
```
Variable morphotactics?

To stay on language, a lame example of variable morphotactics:

(11) a. tóma
   drink.IMP.2SG
   ‘drink (it)!’

   a. tómas
   drink.IMP.2SG
   ‘(you) drink’

   b. tóma-te (un jugo)
   drink.IMP.2SG-1.DAT (a juice)
   drink a juice (for your self)!

   b. te-tómas (un jugo)
   1.DAT-drink.IND.2SG (a juice)
   ‘(you) drink’

   c. tóma-lo (*un jugo)
   drink.IMP.2SG-2.ACC
   ‘drink-it!’

   c. lo-tómas (*un jugo)
   2.ACC-drink.IND.2SG
   ‘(you) drink it’

   d. tóma-te-lo
   drink.IMP.2SG-1.DAT-2.ACC
   ‘drink it up!’

   d. te-lo-tómas
   1.DAT-2.ACC-drink.IND.2SG
   ‘(you) drink it up’

(Baby steps towards Neo Aramaic or Swahili)
Floting morphemes will not work

What could not work, is to try to define “floating” morphemes, which then attach at the right place:

Because it stops being a relational system (and it also probably does not technically work)
What could work

Build complex cells incrementally:
A few issues
A few issues
What could work:

Underspecified relations between cells:

\[
\begin{array}{c}
\text{lo-clitic} \\
\text{PRDM} \\
\text{FTS} \\
\text{PHON} \\
\end{array}
\begin{array}{c}
\text{cell} \\
\text{TAM} \\
\text{SUBJ} \\
\text{ARG-ST} \\
\end{array}
\begin{array}{c}
1 \text{[MOOD indicative]} \\
2 \\
\langle 1 \text{NP, NP} \rangle \\
\end{array}
\begin{array}{c}
\text{cell} \\
\text{TAM} \\
\text{SUBJ} \\
\text{OBJ} \\
\text{ARG-ST} \\
\end{array}
\begin{array}{c}
1 \\
2 \\
\langle 1 \rangle \\
\langle 1 \rangle \\
\end{array}
\begin{array}{c}
\text{PHON} \\
\end{array}
\begin{array}{c}
\langle \sigma_{[o l, N a \lor o]} \rangle \oplus \Box A \\
\end{array}
\]
Concluding remarks

The take home message:

- In order to talk about analogy, we need a proper formalization of analogy.
- I have presented a possible path, others are possible.
- The X-notation is not good enough.

With the model I propose, we can:

- express partial analogies,
- express abstract analogies,
- use underspecification to express general patterns (e.g. ‘morphemes’)
Open questions

Questions

One tricky issue I am aware of:

- Possition classes (e.g. Swahili)

Some important open questions:

- Can all and any generalization be expressed with this system? (i.e. what can analogy not do?)
- How should we represent phonological strings? what is the right level of abstraction? how much phonetics do we want to include?
- How much phonology can we get rid of? (e.g. we can do harmony without phonological process, using underspecification)
- How should we think about redundant constraints?
- What about derivation? (can all derivational cells be listed for a lexeme? would this be required for the model to work?)
That’s it...
## Derivation

<table>
<thead>
<tr>
<th>Verb</th>
<th>INF</th>
<th>1SG</th>
<th>SG</th>
<th>SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to populate'</td>
<td>poblar</td>
<td>pueblo</td>
<td>pueblo</td>
<td>población</td>
</tr>
<tr>
<td>'to tell'</td>
<td>contar</td>
<td>cuento</td>
<td>cuento</td>
<td>'tale'</td>
</tr>
<tr>
<td>'to dream'</td>
<td>soñar</td>
<td>sueño</td>
<td>sueño</td>
<td>'dream'</td>
</tr>
<tr>
<td>'to sing'</td>
<td>cantar</td>
<td>canto</td>
<td>canto</td>
<td>'song'</td>
</tr>
<tr>
<td>'to distinguish'</td>
<td>distinguir</td>
<td>distingo</td>
<td></td>
<td>'distinction'</td>
</tr>
</tbody>
</table>

Note: The above table lists verbs and their corresponding forms in Spanish, including singular and plural forms.
### ATR harmony

**Kasem number system:**

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>yʉkwala</td>
<td>yʉkwalɛ</td>
<td>headscarf</td>
</tr>
<tr>
<td>yukwɛɔrà</td>
<td>yukwɛlì</td>
<td>small skull</td>
</tr>
<tr>
<td>valu</td>
<td>vala</td>
<td>farmer</td>
</tr>
<tr>
<td>vəlu</td>
<td>vələ</td>
<td>traveller</td>
</tr>
<tr>
<td>sugu</td>
<td>sum ~ suni</td>
<td>guinea-fowl</td>
</tr>
<tr>
<td>sʋgʋ</td>
<td>sʋm ~ sʋnɩ</td>
<td>knife</td>
</tr>
<tr>
<td>peeli</td>
<td>peelə</td>
<td>shovel</td>
</tr>
<tr>
<td>pɛɛli</td>
<td>pɛɛla</td>
<td>sardine</td>
</tr>
</tbody>
</table>
lexeme–harmony-attr

\[
\begin{align*}
\text{PAR} \left\{ \begin{array}{l}
\text{PHON} / \begin{bmatrix} \text{vowel} \\ \text{ATR 1} \end{bmatrix} \ldots \begin{bmatrix} \text{vowel} \\ \text{ATR 1} \end{bmatrix} \#/
\end{array} \right\}
\end{align*}
\]

lexeme–class-ɩa

\[
\begin{align*}
\text{PAR} \left\{ \begin{array}{l}
\text{FTS sg} \quad \text{PHON} / X \begin{bmatrix} \text{vowel} \\ \text{HEIGHT high} \\ \text{BACKNESS front} \end{bmatrix} \\
\text{FTS pl} \quad \text{PHON} / X \begin{bmatrix} \text{vowel} \\ \text{HEIGHT low} \\ \text{BACKNESS center} \end{bmatrix}
\end{array} \right\}
\end{align*}
\]