**Motivation & Approach**

- Parallel texts are often less than equivalent due to a range of translation divergences
- Identifying these translation divergences enables fuller utilization of parallel texts in downstream tasks (Vyas et al., 2018; Briakou and Carpuat, 2021)
- The Abstract Meaning Representation (Banarescu et al., 2013) formalism has been expanded to languages other than English in a variety of ways
- We propose an annotation schema to identify specific points of divergence between parallel AMR graphs

  - Present a categorization schema to identify both type and cause of divergence, in order to:
    1. Make the data more adaptable to cross-lingual NLP applications
    2. Identify non-literal translations
    3. Make AMR more cross-lingual consistent (Song et al., 2019)
    4. Provide information on the way in which annotation, semantics, and syntax may play a role in cross-lingual AMR parsing (Damonte, 2019)

- Annotate a small corpus of English-Spanish parallel AMRs (Migueles-Abraira, 2017) from *The Little Prince* with our schema
- Analyze the ways in which cross-lingual parallel AMRs may differ in annotation

**Classification Schema**

- **Structural Divergence**
  - Different Focus
  - Different Non-Core Role Chosen
  - Switch Arg and Non-Core
  - Added/Omitted Arg

- **Cause of Divergence**
  - Semantic Divergence
  - Annotation Divergence
  - Syntactic Divergence

**Annotated Examples**

**No divergence**

- English: Draw me a sheep ...
  - \(d / \text{draw} \cdot 01\)
  - \(:\text{ARG0} (y / \text{you})\)
  - \(:\text{ARG1} (s / \text{sheep})\)
  - \(:\text{ARG2} (i / i)\)
  - \(:\text{mode} \text{imperative})\)
- Spanish: Dibujame una oveja ...
  - Literal translation: Draw me a sheep ...
  - \(\text{dibujar} / \text{draw} \cdot t / \text{tú}\)
  - \(:\text{ARG0} (t / \text{you})\)
  - \(:\text{ARG1} (o / \text{ova})\)
  - \(:\text{ARG2} (y / \text{yo})\)
  - \(:\text{modo} \text{imperativo} )\)

**diffarg synt**

- English: The fourth planet belonged to a businessman.
  - \(\text{belong} / \text{belong} \cdot 01\)
  - \(:\text{ARG0} (p / \text{planet})\)
  - \(:\text{ARG1} (r / \text{businessman})\)
- Spanish: El cuarto planeta era de un hombre de negocios.
  - Literal translation: The fourth planet was of a businessman.
  - \(\p / \text{parecer} )\)
  - \(:\text{ARG0} (e / \text{erupción})\)
  - \(:\text{ARG2} (c / \text{chimenea})\)

**omitnoncore synt (2)**

- English: That is funny!
  - \(f2 / \text{funny}
  - \(:\text{domain} (t2 / \text{that})\)
- Spanish: ¿Qué gracioso!
  - Literal translation: How funny!
  - \(g / \text{gracioso}
  - \(:\text{grade} (t / \text{tan})\)

**focus sem**

- English: Which is your planet?
  - \(p / \text{planet}
  - \(:\text{domain} (a / \text{am-unknown})\)
- Spanish: ¿De qué planeta eres?
  - Literal translation: What planet are you from?
  - \(s / \text{ser-de-91}
  - \(:\text{ARG1} (t / \text{tú})\)
  - \(:\text{ARG2} (p / \text{planeta})\)
  - \(:\text{campo} (a / \text{am-desconocido})\)

**Distribution of Corpus**

- Considerable difference in proportion of structural divergence due to each cause

**Dataset**: https://github.com/shirawein/spanish-english-amr-corpus