EmojiNet: A Machine Readable Emoji Sense Inventory

Sanjaya Wijeratne, Lakshika Balasuriya, Amit Sheth, and Derek Doran

Kno.e.sis Center, Wright State University
Dayton, Ohio, USA
{sanjaya,lakshika,amit,derek}@knoesis.org
http://www.knoesis.org

Abstract. With the rise of social media, 'emoji' have become extremely popular in online communications. People have started using emoji as a new language in social media to add color and whimsiness to their messages. Without rigid semantics attached to them, emoji symbols take on different meanings based on the context of a message. This has resulted in ambiguity in emoji use. Similar to word sense disambiguation, machine readable sense inventories that list emoji meanings are essential for machines to understand emoji without ambiguity. As the first step towards building machines that can understand emoji, this paper presents EmojiNet [1], the first machine readable sense inventory for emoji. It links Unicode emoji representations to their English meanings extracted from the Web, enabling systems to link emoji with their context-specific meaning. EmojiNet is automatically constructed by integrating multiple emoji resources with BabelNet, which is the most comprehensive multilingual sense inventory available to date. The paper discusses its construction, evaluates the automatic resource creation process, and presents a use case where EmojiNet disambiguates emoji usage in tweets. EmojiNet is available online for use at `http://emojinet.knoesis.org`.

Keywords: EmojiNet, Emoji Analysis, Emoji Sense Disambiguation

References

1. Wijeratne, S., Balasuriya, L., Sheth, A., Doran, D.: Emojinet: Building a machine readable sense inventory for emoji. In: 8th International Conference on Social Informatics (SocInfo 2016). Bellevue, WA, USA (11 2016)