We describe the automatic acquisition of a semantic network in which over 7,500 of the most frequently occurring nouns in the English language are linked to their semantically related concepts in the WordNet noun ontology. Relatedness between nouns is discovered automatically from lexical co-occurrence in Wikipedia texts using a novel adaptation of an information theoretic inspired measure. Our algorithm then capitalizes on salient sense clustering among these semantic associates to automatically disambiguate them to their corresponding WordNet noun senses (i.e., concepts). The resultant concept-to-concept associations, stemming from 7,593 target nouns, with 17,104 distinct senses among them, constitute a large-scale semantic network with 208,832 undirected edges between related concepts. Our work can thus be conceived of as augmenting the WordNet noun ontology with related-to links.

The network has been subjected to a variety of evaluative measures, including manual inspection by human judges and quantitative comparison to gold standard data for semantic relatedness measurements. We have also evaluated the network’s performance in an applied setting on a word sense disambiguation (WSD) task in which the network served as a knowledge source for established graph-based spreading activation algorithms, and have shown: a) the network is competitive with WordNet when used as a stand-alone knowledge source for WSD; b) combining our network with WordNet achieves disambiguation results that exceed the performance of either resource individually; and c) our network outperforms a similar resource, WordNet++ (Ponzetto &Navigli, 2010), that has been automatically derived from annotations in the Wikipedia corpus.

Finally, we present a study on human perceptions of relatedness. In our study, we elicited quantitative evaluations of semantic relatedness from human subjects using a variation of the classical methodology that Rubenstein and Goodenough (1965) employed to investigate human perceptions of semantic similarity. Judgments from individual subjects in our study exhibit high average correlation to the elicited relatedness means using leave-one-out sampling (r = 0.77, SD = 0.02, N = 92), although not as high as average human correlation in previous studies of similarity judgments, for which Resnik (1995) established an upper bound of r = 0.90 (SD = 0.07, N = 10). These results suggest that human perceptions of relatedness are less strictly constrained than evaluations of similarity, and establish a clearer expectation for what constitutes human-like performance by a computational measure of semantic relatedness. We also contrast the performance of a variety of similarity and relatedness measures on our dataset to their performance on similarity norms, and introduce our own dataset as a supplementary evaluative standard for relatedness measures.