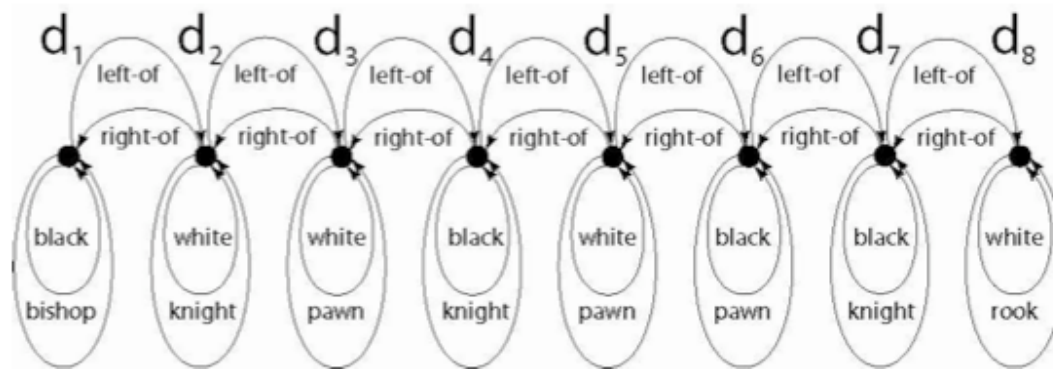
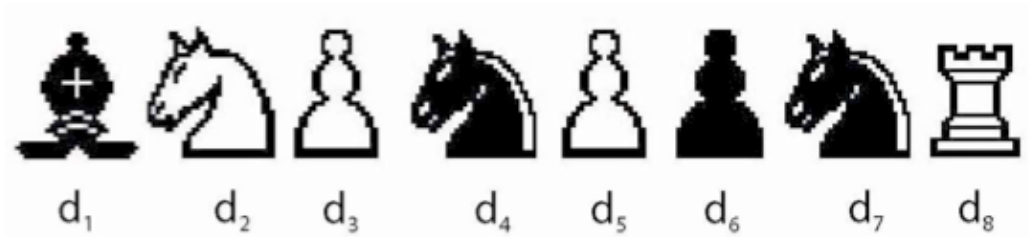


Cost-based attribute selection for GRE [GRAPH-SC / Graph-FP]

Mariët Theune, University of Twente
Pascal Tousef, University of Twente
Jette Viethen, Macquarie University
Emiel Kraahmer, Tilburg University

Graph-based GRE



Cost-based attribute selection

- Graph-based algorithm “ranks” distinguishing descriptions using cost functions.
- For GRE shared task, various cost functions on the TUNA domain were tried.
- **GRAPH-SC**:
 - Costs of properties (edges) are determined by their frequency.
 - For each edge e : $\text{Cost}(e) = -\log_2(P(e))$.
 - The more frequently a property occurs in the corpus, the cheaper it is.
- **GRAPH-FP**: A domain specific extension of GRAPH-SC, where frequently used (potentially redundant) properties can be added at zero costs:
 - **Furniture**: colour
 - **People**: hasBeard and hasGlasses

Results

	Furniture domain		People domain	
System	Dice	PRP	Dice	PRP
GRAPH-SC	0.66	20.0	0.66	22.1
GRAPH-FP	0.71	30.0	0.67	25.0

Discussion

- In Gatt et al. (2007) the (best version of the) **Incremental Algorithm** has a somewhat lower mean Dice score (0.64), and a substantially lower PRP (10%).
- **Lots of room for improvement:**
 - Different speakers in the corpus seem to have different preferences. Would be interesting to try “speaker-dependent cost functions”.
 - Our approach does not deal properly with interdependencies between properties (e.g., hairColour and hasHair).
 - We intend to rework the algorithm such that it does not output a single best description, but a ranked list of possibilities. (Evaluate with the Mean Reciprocal Rank metric?)