Mechanism Design for Ontology Alignment

Piotr Krysta, University of Liverpool (P.Krysta@liverpool.ac.uk)
Minming Li, City University of Hong Kong, China (minming.li@cityu.edu.hk)
Terry R. Payne, University of Liverpool (T.R.Payne@liverpool.ac.uk)
Nan Zhi, University of Liverpool (N.Zhi@liverpool.ac.uk)

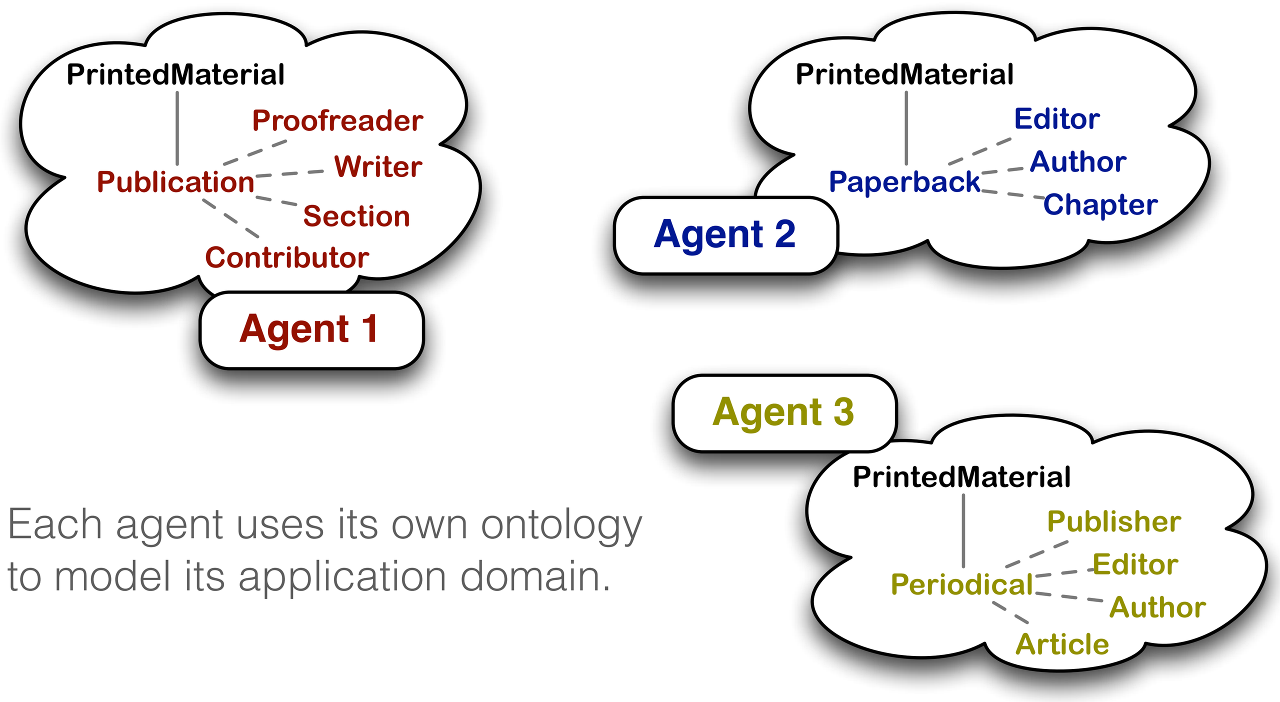
Abstract: The aim of the ontology alignment problem is to find meaningful correspondences between two ontologies represented as collections of entities. This problem can be modelled as a novel mechanism design problem on an edge-weighted bipartite graph, where each side of the graph holds each agent's private entities, and the objective is to maximise the agents' social welfare. Having studied implementation in dominant strategies with and without payments, we report on findings that for truthful mechanisms, these problems need to be solved optimally. We also study greedy allocation rules with a first-price payment rule, and implementation in pure, mixed & Bayesian Nash equilibria, and have found tight bounds on the price of anarchy and stability.

Motivation and Setting

Within open agent environments, agents may differ in the way they model their domain (i.e. they assume different ontologies that are *semantically heterogeneous* for a given shared domain). To enable meaningful communication, the agents need to *align their ontologies* (i.e. map concepts from one ontology to the corresponding ones in another ontology). However, the choice of correspondences will depend on the goals of each agent, and thus each agent may have different preferences ω over the choice of correspondences. Hence, a mechanism is needed to find the right mappings given the individual preferences of each agent.

Agents *may have access to various different correspondences* (from shared repositories or garnered from previous transactions with other agents), some of which may be irrelevant to the current communication, or may result in ambiguity (i.e. resulting from a one-to-many mapping). Thus, the aim here is to find a *matching* (i.e. an injective alignment) based on a set of individual bids (put forward by each agent) for candidate correspondences.

The setting is modelled as an edge weighted bipartite graph, where the nodes represent the concepts in each agents ontology, and the edges represent the candidate correspondences. However, agents may not wish to reveal what correspondences they know, or how important they are (i.e. their weight). We explore this scenario from a Mechanism Design perspective.



Each agent uses its own ontology to model its application domain.

⟨Publication ⟷ Paperback, ω⟩
⟨Publication ⟷ Periodical, ω⟩
⟨Proofreader ⟷ Publisher, ω⟩
⟨Proofreader ⟷ Editor, ω⟩
⟨Writer ⟷ Editor, ω⟩
⟨Writer ⟷ Author, ω⟩
⟨Contributor ⟷ Author, ω⟩
⟨Section ⟷ Chapter, ω⟩
⟨Section ⟷ Article, ω⟩
...

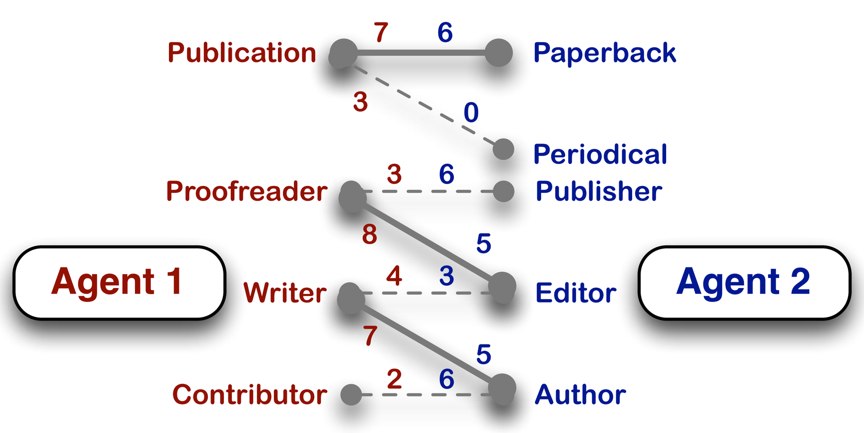
To support communication, each agent has access to various correspondences that map between concepts, and has to bid for the one believed to be the most relevant.

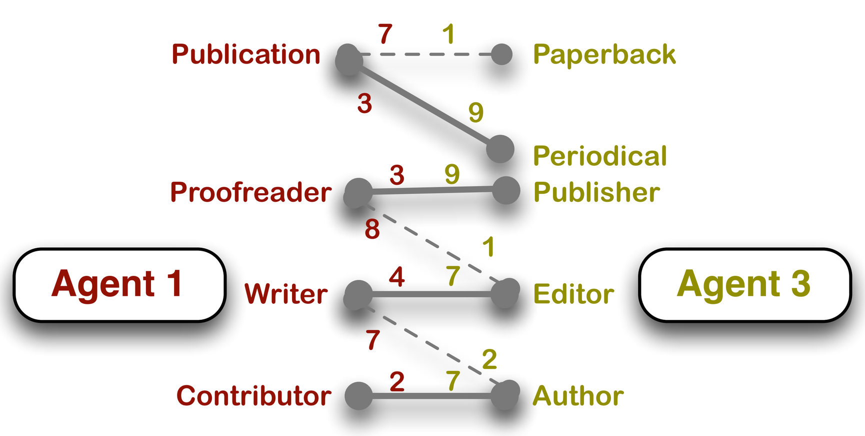
Candidate Correspondences
Note - ω refers to the weight, which may be private to each agent

Mechanism Design with Payment

This is modelled on the assumption that agents incur a cost in revealing their true value, but will gain if they achieve their desired alignment. To explore truthful mechanisms, we can use the classic VCG mechanism with Clarke payment if we are willing to solve the problem optimally. The weight of each correspondence is calculated by simply adding the individual bids for that correspondence. The resulting matching can be found using a greedy (approximation) algorithm. By analysing this approach, we found that ***it is not possible to have a faster, non-optimal, approximate and truthful mechanism for our problem.***

Based on past transactions, Agent 1 prefers alignments that map **Publication** to the concept **Paperback**, and thus weights the correspondences that are successful when transacting with others interested in Paperbacks





Agent 2 is interested in **Paperbacks**, and knows about the concepts **Publisher**, **Editor** and **Author**, but is ambivalent as to how these map to other concepts.

Agent 3 however, is specifically interested in **Periodicals**, and has a preference over certain correspondences, that lead to a different matching.

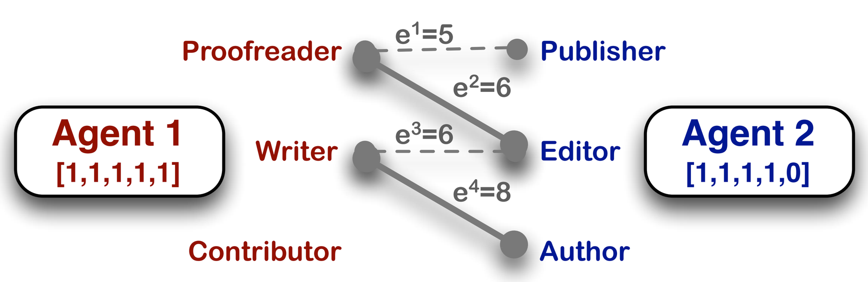
Theorem 1. For the alignment problem with payment, any mechanism not adopting an optimal solution when agents declare their true valuations is either non-truthful, or has an approximation ratio of at least 2.

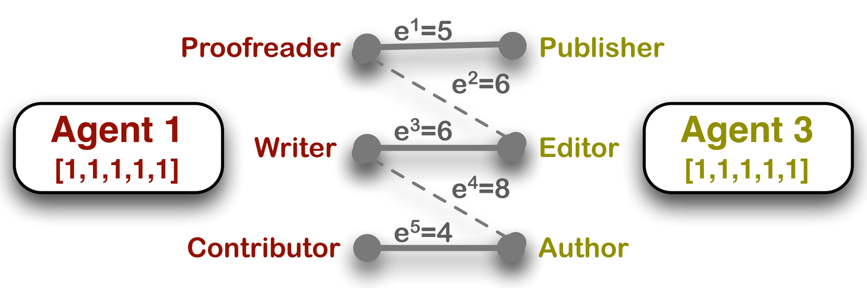
Theorem 2. For the alignment problem with payment, any deterministic mechanism which does not adopt optimal solution when agents declare true valuation, is either non-truthful, or is maximal-in-range.

Mechanism Design without Payment

A mechanism setting where the value of each correspondence was considered public (and thus agents cannot misrepresent their valuations) was also studied. In this setting, agents submit a boolean vector identifying which of the correspondences are preferable. Edges that are selected by both agents are considered with their public value. However, if an edge is not selected by at least one agent, then its value will be zero. A matching is then determined from these edges by maximising social welfare. ***The aim is to incentivise both agents to declare their choice of edges (i.e. correspondences) truthfully***.

In the non-payment setting, Agent 1 & 3 both express an interest in all 5 correspondences. Thus, the matching $\{e^1, e^3, e^5\}$ is found that maximises social welfare of 15.





Agent 2 is not interested in correspondences mapping to **contributor**, and only expresses interest into edges 1-4. Despite agent 1 preferring all edges, only the first four are considered. Therefore, the matching $\{e^2, e^4\}$ maximises social welfare with a value of 14.

A polynomial time algorithm was designed which determines that, given an instance, if a deterministic truthful mechanism exists with a bounded approximation ratio; if so, then the optimal solution is found. However, if such a mechanism does not exist for the bid, then we show there is no truthful mechanism with a bounded approximation ratio.

Theorem 3. There are no randomized mechanisms that are universally truthful and have approximation ratios better than 2 for the setting.

Theorem 4. There are no randomised mechanisms that are truthful in expectation and have approximation ratios better than 1.333 for the setting.

UNIVERSITY OF LIVERPOOL

CityU 香港城市大學 City University of Hong Kong 專業 創新 胸懷全球 Professional · Creative For The World

PRINTED BY SCIENCEPOSTERS.co.uk