Report: 1st ESSENCE Autumn School

Pavlos Andreadis

November 02, 2014.

The 1st ESSENCE Autumn School was an interdisciplinary set of 5 tutorials and 3 lectures, as well as a course on managing your research project and an open talk on doing a PhD. The school centered around the theme of evolving semantics, with tutorials examining issues such as

- theory and problem classes in ontology matching,
- how iterated learning leads to structure in language,
- dialogue semantics and pragmatics,
- an overview of reinforcement learning and communication in mutiagent systems,
- and argumentation in dialogue.

The workshop was also a good opportunity for networking, and functioned as a kick-starter for at least one independent project. The attendees were a diverse group and, except for ESSENCE fellows, also included post-graduate students in semantics, linguistics, decision theory, knowledge bases (Law), as well as at least one student representing a start-up with a related business goal.

ESSENCE is a Marie Curie Initial Training Network, involving universities and other goverment and industrial institutions across Europe. Universities involved are in Edinburgh, Trento, Amsterdam, Barcelona, and Brussels. Its acronym stands for Evolution of Shared SEmaNtics in Computational Environments and its research revolves around the emergence and evolution of meaning through intelligent agent communication. For more information on ESSENCE you can visit `https://www.essence-network.com`.

1 Day 1, Monday 27th November - Tutorials 1&2

The school opened with a talk from University of Edinburgh's Michael Rovatsos, coordinator of the ESSENCE project. After giving the financial and geographic scope of the project, Rovatsos proceeded to present the main observation on which the project is structured, namely:

- People are good at resolving communication conflicts: they are robust, resilient, flexible, and adaptable.

ESSENCE's goal is to translate these enabling capabilities to machines. In doing so, it focuses on the computational aspect while also building the necessary theoretical foundations. Their main approaches are:

- Modelling meaning in communication as an evolving process of negotiation.
- Understanding the foundations of representation, reasoning, decision making, and learning in communication.
- Developing computational models of evolving semantics in human and artificial collectives.

After some analysis on these points and a few closing remarks on the workshop's organization we were on to our first tutorial!

1.1 Tutorial 1 - Pavel Shvaiko: *Ontology Matching*

Expecting a diverse audience, Pavel Shvaiko, from Informatica Trentina, attempted an informal introduction to the notions relating to *Ontology Matching* (OM). He described the *Semantic Web* as, not one or x people's common ontology, but as many people's ontologies, "a meaningful mess", with heterogeneity "giving rise to features, not bugs". The goal is to reconcile mismatches between formalised knowledge that results from differently expressed resources. Such mismatches can occur because of different languages and terminologies, modelling procedures, granularity and perspective, scope, or application. The problem of OM can be seen as that of reducing heterogeneity by providing an alignment between two ontologies (all examples and descriptions mentioned two ontologies) via a matching procedure, and then using the results to generate a process to merge/transform information from one to the other.

An allignment will include correspondences between ontological entities. A correspondence e is a tuple $< e, e', r >$ indicating the relation r between entities e and e' ($e\, r\, e'$). In the context of ontology matching, a mapping will generally refer to a directed alignment, though this is far from the norm, with mapping sometimes referring to the process itself.

A large portion of the first part of the tutorial was dedicated to explaining the differences between specific cases of OM along with broad examples. A few applications are query answering, linked-data interlinking, peer-to-peer information sharing, and web service composition. The second part summarised a variety of approaches to OM, in particular: String-based, Language-based, Linguistic resources, Constraint-based, Extensional, Tree-based, Graph-based, Probabilistic, Model-based, and Context-based matching, while looking into issues such as social and collabrative OM, and state of the art systems. Shvaiko closed with a mention to his new book "Ontology Matching", published by Springer.

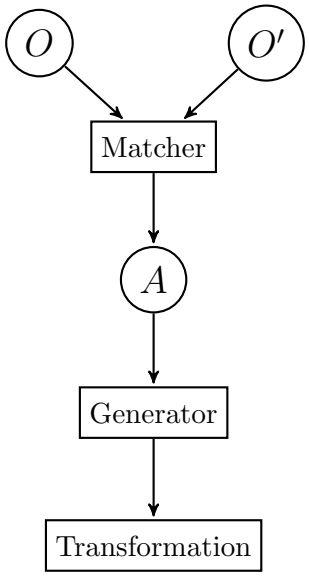

Figure 1: A basic representation of ontology matching.

1.2 Tutorial 2 - Simon Kirby: *Iterated Learning*

Simon Kirby, from the University of Edinburgh, began by introducing grammar as a systematic structure, unique to language, and urged us to wonder on the cause of this. Foreshadowing his analysis, he explained this as the result of a trade-off between two partially competing pressures: representational compressibility and expressivity.

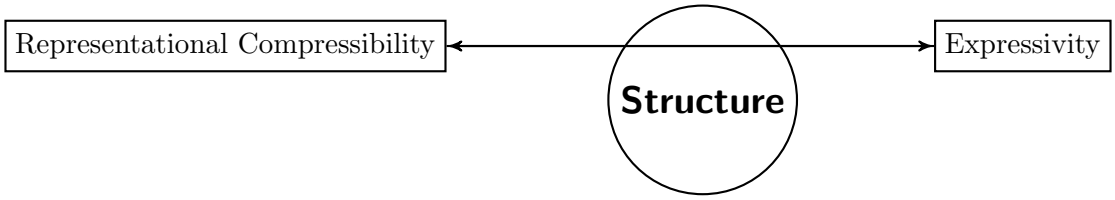

Figure 2: The trade-off between representational compressibility and expressivity gives rise to linguistic structure.

Gene selection is mentioned in the evolution of language but it is cultural transmission that is most emphasized. "We are not born with a complete language." and "Two complex adaptive systems interact." are key phrases here.

The rest of the tutorial was focussed on investigating how iterated learning leads to structure in language. The presentation outline was as follows:

- Emerging structure in a psychological lab experiment.
- Simple, agent-based simulation replicating experiment.
- Another lab experiment showing the emergence of symbols.

- A quick overview of current work in various domains and species.

The attendees were left with the question: "Should understanding the origins of structured behaviour change the way in which we construct intelligent systems?"

A highlight of the talk was an experiment showing the evolution of a monkey culture so that it used *Tetrominoes* (yes, those blocks in Tetris!) when solving a particular task. Kirby concluded with two points:

- Cultural transmission of sets of behaviours leads to systematic structure, and
- Whenever there is iterated learning of behavioural repertoires, we should find compressible, systematic structure emerging.

2 Day 2, Tuesday 28th November - Lectures, Tutorial 3 & After-dinner speech

Day 2 was information intense, featuring three lectures on current research, the third tutorial, and an open talk on doing your PhD.

2.1 Lecture 1 - Vincenzo Maltese: *Linguistic and Knowledge Resources*

After a reintroduction to the problem of semantic heterogeneity, and use cases such as semantic matching, Vincenzo Maltese, representing the KnowDive Research Group from the University of Trento, proceeded to give two timelines of available resources. Starting with WordNet and MultiWordNet (2002), he explained the limitations of such approaches:

- Nodes in similar positions do not share ontological properties;
- Classes exhibit space and time bias;
- Some concepts are similar in meaning;
- Some concepts are actually individuals (as in, King Robert the 4th).

Maltese then presented a second timeline of knowledge resources, specifically: the CYC ontology (1984), the SUMO ontology (2002), the YAGO ontology (2008), Freebase (2010), and the Linked Data Cloud (since 2007). He concluded by decsribing Trento University's approach to creating a knowledge resource: A WordNet-free version with more ontologies and less linguistic resources by turning the hierarchy in WordNet into an ontology, all based on an entity-centric view of the world. An important result of this approach is that of removing language bias (usually bias for the English language) since terms and ideas connect with each other only through abstract middle nodes. When an idea is not existent in a language, there is an attempt to replace it with a description (e.g. there is no word for *harbour* in Mongolian).

2.2 Lecture 2 - Carles Sierra: *Agreement Computing*

Carles Sierra, from the Artificial Intelligence Research Institute (Spanish Research Council), proposed an alternative programming paradigm in *Agreement Computing*. Essentially, he argues, software should be built around explicit expectations of its behaviour by the user, through understanding human interaction. Programs are described as "ever-adapting norm solvers". The approach does not differentiate between human and artificial users/agents.

The online game *Diplomacy* was presented as a case study, with empasis on the role of negotiation for agreement forming. A subcase focuses on aiding humans in reaching agreement between each other. Sierra concludes by mentioning the need for simpler norm descriptions as well as for understanding the evolution of communities, while pointing to the website IFTTT, as a good first paradigm.

2.3 Lecture 3 - Alan Bundy: *Reformation: A Domain-Independent Algorithm for Theory Repair*

The third and final lecture was given by Alan Bundy, from the University of Edinburgh. Alan Bundy is part of the Galileo project, which aims at finding incosistencies in threories given experimental results. He presented the algorithm *Reformation*, an evolution of the *Standard Unification Algorithm* [Baader and Snyder, 2001] for *Theory Repair*. Reformation is a heuristic's based algorithm for correcting contradicting semantics.

Bundy makes the case for automatic repair of semantics by pointing to past experience, where a stated minor error percentage in the database translated to millions of preexisting errors. Incosistencies in theory can often be explained by the presence of an unobserved phenomenon (e.g. latent heat, dark matter; note the adjectives). Near the end, he emphasizes the need for heuristics when pruning the huge search space of such problems.

2.4 Tutorial 3 - David Schlangen: *Dialogue Semantics and Pragmatics*

Closing the 2nd day, David Schlangen, from the University of Bielefeld, presented an analysis over shared meaning and shared content creation, giving the basics and computational models. The presentation was grounded on a simple example of two agents involved in dialogue. In this, agent Alice tells agent Bob that "There is a tiger behind those bushes!" In a simplified world of two objects, obj_{tiger} and $obj_{chicken}$, and two signals, *tiger* and *chicken*, there are two ways of defining a language that discriminates between the two situations, namely:

- $tiger \rightarrow obj_{tiger}$, $chicken \rightarrow obj_{chicken}$, and
- $tiger \rightarrow obj_{chicken}$, $chicken \rightarrow obj_{tiger}$.

Essentially, there is no inherent property, either of signal or object, that defines the correct method of referring to these objects in this coordination problem. So what helps us decide on how to refer to each object?

Our coordination devices are explicit agreement and convention, as well as precedence and salience. The latter two were beautifully hammered in when Schlangen asked the participants to provide the same answers to a set of questions without prior coordination. Apparently we humans are quite good at deciding on meeting points and specific hours on the fly.

Agents in dialogue perceive, decide, and act, thus closing the loop. While doing so, they maintain a mental state of beliefs, beliefs about the other agent's beliefs, beliefs about the other agent's belief's about its beliefs and so on. Mutual beliefs are impossible to maintain, given that the two agents don't share a single mind. Essentially, there can be no certainty of alignment, a fact exemplified by the prolonged information acknowledgement procedures utilised in flight control communication.

Some of the issues mentioned were conventional and non-conventional meaning, communicative intentions, timing and turn-taking in conversation. The first part of the talk included a Nao experiment, aimed at creating a benchmark for realistic turn-taking in human conversation. It is interesting to note, that we do not simply wait for the other speaker to finish before initiating our response. Mechanisms considered in explaining this included vocal pitch monitoring and identification of coherent arguments/sentences.

Part 2 of the tutorial focused on computational models, outlining the approaches of

- Finite State Analysis (FSA) [Oviatt, 1994; Mclean, 1998],
- Belief-Desire-Intention (BDI) [Allen, 1995; Perrault and Allen, 1980],
- Conversational scoreboard [Larsson and Traum, 2000],
- and POMDP [Singh et al., 2000; Williams and Young, 2007].

Schlangen proceeded to present *IU*, a temporally fine-grained model, as well as the NUMBERS and PENTO-10 systems. Both impressive voice analysis systems with realistic, fluid interaction between user and system. The 2013 PENTO system (which is open source) integrates human gaze and gestures.

2.5 After-dinner speech - Carles Sierra: *How to do the PhD and survive it*

Day 2 ended with an informal talk by Carles Sierra on how to approach a carreer as a researcher and what problems we will face in these early stages. A truly insightful analysis built around Santiago Ramony Cajal (1899)'s book "Advice for a young investigator".

3 Day 3, Wednesday 29th October - Tutorials 4&5

3.1 Tutorial 4 - Matthijs Spaan: *Decision-theoretic Approaches in Multiagent Systems*

Matthijs Spaan's, from the Delft University of Technology, lecture was structured into two parts: The first gave an overview of classic planning, MDPs, and POMDPs, including a description of various POMDP solvers and an explanation of the piece-wise

linearity and convexivity of the value function. The second part dealt with multiagent (MA) planning, rating different Reinforcement Learning (RL) models on the axes of agent communication and observability. There was mention of the inherent trade-off of coordination and local information exploitation, as well as Spaan et al. (2006)'s paper on DEC-POMDPs.

Spaan called for the need of handling heterogenous and a priori unknown semantics within the context of MA planning, an area which he considers ripe for research. He argues, that the key to reducing the complexity of MA systems lies in factorising the policy space in accordance to offline communication policies, that take advantage of local interactions and exploit the domain structure. He continues briefly with event-driven models, real-time execution strategies and generalized Semi-Markov Decision Processes before concluding with the related challenges of structure indentification and considering richer communication. Spaan pointed out the book *Reinforcement Learning: State of the Art* [Wieving and Ottelo eds, 2012] while extending an invitation for contacting him on information regarding RL solvers (also see http://masplan.org).

In our break discussion, Spaan mentioned an untackled RL research direction which I understood as *execution-time translation of input signals into RL models.* An automation of the RL modelling procedure, such that it can be applied online, perhaps using an ontological description of the world.

3.2 Tutorial 5 - Sanjay Modgil: *Argumentation and Dialogue*

The school's last tutorial was given by Sanjay Modgil, of King's College London. Given the guarantee of it being the same tutorial that was given in *EASSS 2014 - 16th European Summer School*, which I had attended, I decided to do some coding instead. However, this is by far the best tutorial I have attended, a sentiment that seems to be shared by quite a few people. On that note I will only give a brief outline and sample of the talk. You can find the slides for this presentation on:`http://www.dcs.kcl.ac.uk/staff/smodgil/SlideDownload.zip`.

Modgil's talk is divided into two parts. The first one titled *models for agent reasoning and communication*, teaches the fundumentals of reasoning using logic-based argumentation. It starts by giving *Dung's Abstract Argumentation Theory.* A Dung Argumentation Framework (AF) is a directed graph $(Args, Att)$ where the nodes $Args$ denote arguments, and Att denotes a binary relation where one argument *attacks* the other in an attempt to have it *rejected* [Dung, 1995]. Given a set of well-formed formulas (wff) such as:

$$\begin{aligned}\Delta = \{ & q :- p, not\ s; (\text{i.e. } q \text{ is true if } p \text{ is true and } s \text{ isn't}) \\ & s :- not\ g; \\ & g :- m; \\ & p; \\ & m\}\end{aligned}$$

we can construct a logic programming instantiation of a Dung AF. Then, according to the selected protocols of argumentation, we can define subsets of arguments that exhibit some properties, and in so doing declare the winning set of argument(s). In this example, an argument could be:

$$X = [q \ :- \ p, not\ s;\ p],$$
$$Y = [s \ :- \ not\ g], or$$
$$Z = [g : -m;\ m].$$

Here, q, s, and g are the claims of arguments X, Y, and Z, respectively. Now, if you were to be presented with the argument X, given the set Δ, which is a predefined mutually accepted *truth*, you can attack it with, say, Y. The other agent could then respond, however, by attacking Y with Z. More descriptively, her claim of q does not hold if s is true, which Y claims, only to be shown incorrect by Z, which states that since g holds, s cannot hold.

An argument is *justified* if it is not succesfully attacked, i.e. if any argument attacking it has been succesfully attacked. Arguments can be either *justified*, *rejected*, or *undecided*. We say that an argument Z that attacks an argument Y attacking an argument X, *defends* or *reinstates* X. We say that a set of arguments S is *admissible*, if it is *conflict free* and all its contained arguments are defended against attacks. S is *conflict free* if no argument in the set attacks another in the set.

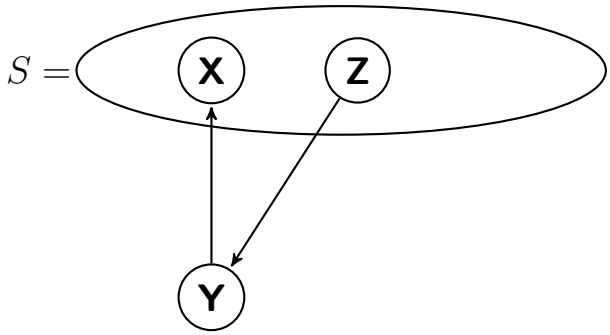

Figure 3: S is a conflict free extension.

Part 1 continues describing this *Dung's calculus of opposition* and related semantics, such as complete and preferred extension, while giving many examples for the audience to work with. The issue of handling multiple justified *extensions* (i.e. sets of arguments) is well analysed while a particulary interesting point involves argumentation with preferences. Modgil explains how argumentation is used in distributed non-monotonic reasoning (dialogue), explains *rationality postulates* (properties that have to hold for arguments contained in a *complete* extension) and presents the ASPIC+, a framework of intermediate abstraction between instantiating logic and an abstract argumentation framework [Modgil and Prakken, 2013].

Part 2 of Modgil's talk is titled *argumentation based dialogue* and gives argument game proof theories [Modgil and Caminada, 2009] for deciding whether an argument is

in a preferred or the grounded extension. This part focuses, essentially, on analysing different sets of rules on how to play these games and how to declare a winner. It then proceeds to model dialogues in this light, where the communication language defines legal *locutions*, consisting of *speech acts* and *content* (e.g claim(a = "Helicopters are cooler than cars."), why(a), argue(a since b and $b \rightarrow a$)). The tutorial concludes with argumentation enabled agent architectures and *schemes*, which are generic templates for arguments with associated critical questions.

4 Day 4, Thursday 30th October - Course

4.1 Course - Sara Shinton: *Taking Control of your Research Project*

The 1st ESSENCE Autumn School concluded with an interactive, in-depth analysis of methodologies and approaches to planning your research project, and then staying on course. With Sara Shinton, from Shinton Consulting Ltd, we focussed on issues such as *who the project stakeholders are*, *the importance of networking*, and creating a mind-map and objectives tree for your research.

5 Conclusion

To conclude, I found the school to be an amazing experience and definetely suggest attending future workshops, given a minimal overlap with your interests. I will be happy to share further details on anything that caught your attention.