Hypernetwork-based peer marking for scalable certificated mass education

Conference or Workshop Item

Hypernetwork-based Peer Marking for Scalable Certificated Mass Education

Jeff Johnson, Cristian Jimenez-Romero, David Rodrigues, Jane Bromley and Alistair Willis
The Open University, Milton Keynes, MK7 6AA, UK

In the context of the need for massive free education for the Complex Systems Society and the UNESCO Complex Systems Digital Campus, scalable methods are essential for assessing tens of thousands of students' work for certification[1]. Automated marking is a partial solution but has many drawbacks. Peer marking, where students mark each others' assignments, is a scalable solution since every extra student is an extra marker. However there are concerns about the quality of peer marking, since some students may not be competent to mark the work of others. Some students are better than others and often the best students are well qualified to assess the work of their peers. To make peer marking high quality we are using new hypernetwork-based methods to extend previous methods[2] to discover which students are good markers and which students are less good as a course progresses.

Peer marking is becoming increasingly used in education. It has the obvious pedagogic advantage that marking other students' assignments gives students insights into how well or otherwise the marker themself performed. This alone makes peer marking attractive. To allow for variable quality in marking it is common for students to mark the assignments of two or three other students, but marking higher numbers becomes onerous and is rarer. The simplest approach is to average the peer-marks, but this is crude and unreliable. We seek to be able to discriminate good markers from bad markers, and to assign reputations to them so that poor markers can be excluded from the final score given to an assignment, which is essential when the mark counts towards certification.

The basic idea is that good markers will be more highly connected than bad markers. Suppose a class of students have ten assessment tasks, with each student producing an answer for each task. Consider two students peer-marking exactly the same set of answers. If the are good markers, their marks will tend to be similar on each question, being close to the 'correct mark', and they will be *10-similar*. Now consider two poor students marking exactly the same set of ten answers. These students will give marks that deviate a lot from the 'correct' score and are unlikely to give the same marks, *e.g.* they may be arbitrarily higher and lower. These poor students are likely to be relatively disconnected. Thus being highly connected is *necessary* for two markers to be good while being relatively disconnected is *sufficient* for one or both of a pair of markers to be bad.

In our experiment we have fifty Open University PhD students and three researchers studying a specially prepared short course on Global Systems Science. For each lesson, these students read a short text and complete an assignment with five questions. They upload these to our Étoile peer marking platform. The students peer-mark the assignments of three other students, and then they mark their own assignment. If the self-assessment is different from the peer-assessment this gives useful diagnostic information. An important feature of our experimental design is that we make the peer-marking symmetric – if student *a* marks student *b*'s answer then student *b* marks student *a*'s answer. A consequence of this is that for each assignment the students are assembled into groups of four, with each student marking the work of the other three in the group. Such groups are hypersimplices[3], $< a, b, c, d; R >$, where *a*, *b*, *c*, and *d* are students and *R* is the 4-ary relation that binds them together. Clearly *R* is a very interesting relation that give a lot of diagnostic information. For the next assignment the students are grouped differently. For example, as $< a, b, e, f\,; R' >$ and $< c, d, h, g\,; R' >$. We have seven assignments, so each student is related to 7 x 4 = 28 answers. By designing the hypernetwork of hypersimplices appropriately, for example, eight student markers can pairwise share 12 answers. Thus two good students can have up to twelve similar marks and be 12-similar, while bad markers will be less highly connected. We are investigating the hypernetwork connectivity of marking groups of eight and sixteen students. Our experiment will investigate various underlying topologies for establishing marker reputations in a robust way, to provide demonstrably high-quality peer marking.

The experiment will be complete at the end of May 2014. Our full paper will publish the data and report the results of the hypernetwork-based method of identifying good and bad peer marking.

---

[1] Johnson, J.H., Willis, A., Hales, D., Louçã, J., Bourgine, P., Kolhase., Étoile Cascades Ideas', European Conference on Complex Systems, ECCS'11, Vienna, September 2011.

[2] Jimenez-Romero, C., Johnson, J., De Castro, T., 'Machine and social intelligent peer-assessment systems for assessing large student populations in massive open online education', 12th European Conference on e-Learning ECEL-2013, SKEMA Business School, Sophia Antipolis, France, 30-31 October 2013

[3] Johnson, J. H., *Hypernetworks in the science of complex systems*, Imperial College Press, (London) 2013.