EXTREMAL UNKNOWNS

ROBERT SANDERS

Identifying key missing information is an important part of the survey process. Complexity Zoology has the capacity to identify a subset of unknown inclusions that are sufficient to decide all other unknowns. To be precise, recall that an inclusion $\mathsf{C}_1 \subseteq_W \mathsf{C}_2$ can exist in one of three possible states: proven, disproven, or open. However, since Complexity Zoology's data set is incomplete, it is possible that the system cannot infer which of the three possibilities applies to a given inclusion. When this occurs, we say that the inclusion is *unknown*, not to be confused with the inclusion being open. In short, to say that an inclusion is open is to say that it does not have a known proof or disproof, while to say that an inclusion is unknown means that Complexity Zoology itself does not know the status. We denote an unknown inclusion by $\mathsf{C}_1 \overset{?}{\subseteq}_W \mathsf{C}_2$.

Ideally, we would like to complete the data set so that there are no unknown inclusions. However, this is a process that potentially involves a great deal of redundancy, since listing a formerly unknown inclusion as proven (for example) may decide several other unknown inclusions. Therefore, it is helpful to identify a subset of unknown inclusions that are sufficient to decide all others. We refer to these as *extremal unknowns*.

A *most likely extremal unknown* is, roughly, an unknown inclusion that, if it were listed as proven, would not result in any other inclusions being listed as proven. Complexity Zoology considers extremal unknowns separately for each of the worlds, and for simplicity inference rules involving multiple worlds or involving operatros are not considered. Thus, the only inference rules that are relevant to determining whether Complexity Zoology considers an unknown inclusion to be a most likely extremal are

$$p(\mathsf{C}_1 \subseteq_W \mathsf{C}_2) \mathbin{\&} p(\mathsf{C}_2 \subseteq_{W^*} \mathsf{C}_3) \Longrightarrow p(\mathsf{C}_1 \subseteq_W \mathsf{C}_3),$$
$$p(\mathsf{C}_1 \subseteq_W \mathsf{C}_2) \mathbin{\&} p(\mathsf{C}_2 \subseteq_{W^*} \mathsf{C}_3) \Longrightarrow p(\mathsf{C}_1 \subseteq_W \mathsf{C}_3),$$

where W is the world under consideration and W^* is the transitive dual. In testing whether an unknown inclusion $\mathsf{C}_1 \overset{?}{\subseteq}_W \mathsf{C}_2$ is a most likely extremal, Complexity Zoology tests whether there exists a third class C_3 that is distinct from C_1 and C_2 in W satisfying one of the following two conditions:

$$\mathsf{C}_1 \overset{?}{\subseteq}_W \mathsf{C}_3 \mathbin{\&} p(\mathsf{C}_2 \subseteq_{W^*} \mathsf{C}_3),$$
$$\mathsf{C}_3 \overset{?}{\subseteq}_W \mathsf{C}_2 \mathbin{\&} p(\mathsf{C}_3 \subseteq_{W^*} \mathsf{C}_1).$$

If such a C_3 exists, then the system concludes that the unknown is *not* most likely extremal.

Similarly, a *least likely extremal unknown* is intended to be an unknown inclusion that, if listed as disproven, would not result in any other unknown inclusions being decided as disproven. In this case, the relevant inference rules are

$$d(\mathsf{C}_1 \subseteq_W \mathsf{C}_2) \mathbin{\&} p(\mathsf{C}_3 \subseteq_{W^*} \mathsf{C}_2) \Longrightarrow p(\mathsf{C}_1 \subseteq_W \mathsf{C}_3),$$
$$d(\mathsf{C}_1 \subseteq_W \mathsf{C}_2) \mathbin{\&} p(\mathsf{C}_1 \subseteq_{W^*} \mathsf{C}_3) \Longrightarrow p(\mathsf{C}_3 \subseteq_W \mathsf{C}_2).$$

Complexity Zoology checks whether a given unknown inclusion $\mathsf{C}_1 \overset{?}{\subseteq}_W \mathsf{C}_2$ is a least likely extremal by checking whether there exists a class C_3 that is distinct from C_1 and C_2 in W and satisfies at least one of the following:

$$\mathsf{C}_1 \overset{?}{\subseteq}_W \mathsf{C}_3 \mathbin{\&} p(\mathsf{C}_3 \subseteq_{W^*} \mathsf{C}_2),$$

$$\mathsf{C}_3 \overset{?}{\subseteq}_W \mathsf{C}_2 \mathbin{\&} p(\mathsf{C}_1 \subseteq_{W^*} \mathsf{C}_3).$$

Complexity Zoology lists extremal unknowns of both types for each world. It also specifies whether each extremal is provable, disprovable, or completely unknown (neither provable nor disprovable).

Proposition. *Deciding the extremal unknowns is sufficient to decide all the unknowns for a world.*

The unknowns that Complexity Zoology has identified as extremal have resulted in several interesting questions. They have also been extremely useful in filling in large portions of the data very quickly, in particular, a disproof for a most likely inclusion or a proof for a least likely inclusion tend to settle many other unknowns.