

# Machine Learning a Probabilistic Network of Ecological Interactions

Alireza Tamaddoni-Nezhad<sup>1</sup>, David Bohan<sup>2</sup>,  
Alan Raybould<sup>3</sup>, and Stephen Muggleton<sup>1</sup>

<sup>1</sup> Department of Computing, Imperial College London, London, SW7 2AZ, UK  
Email: {a.tamaddoni-nezhad,s.muggleton}@imperial.ac.uk

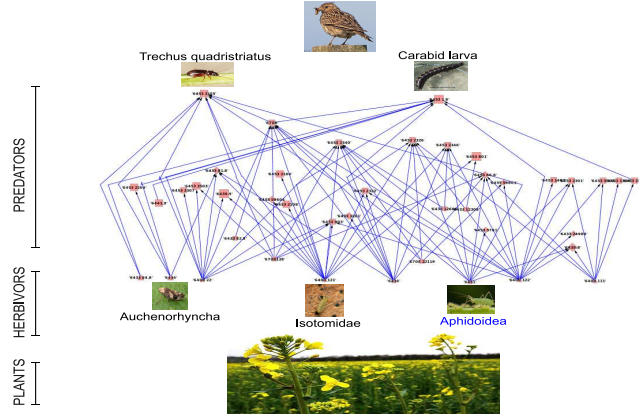
<sup>2</sup> Rothamsted Research, West Common, Harpenden, Herts, AL5 2JQ, UK  
Email: david.bohan@rothamsted.ac.uk

<sup>3</sup> Syngenta Ltd, Bracknell, Berkshire, RG42 6EY, UK  
Email: alan.raybould@syngenta.com

**Abstract.** In this paper we demonstrate that Abductive ILP can generate plausible and testable food webs from ecological data. In this approach, unlike previous applications, the abductive predicate ‘eats’ is entirely undefined before the start of the learning. We also explore a new approach for estimating probabilities for hypothetical ‘eats’ facts based on their frequency of occurrence when randomly sampling the hypothesis space. The results of cross-validation tests suggest that the trophic networks with probabilities have higher predictive accuracies compared to the networks without probabilities. The proposed trophic networks have been examined by domain experts and comparison with the literature shows that many of the links are corroborated by the literature. In particular, links ascribed with high probability are shown to correspond well with those having multiple references in the literature. In some cases novel high probability links are suggested, which could be tested. [EXTEND ABSTRACT WITH RESULTS ON FUNC. GROUPS]

## 1 Introduction

Machine Learning has the potential to address many challenging problems in ecological sciences [4]. Discovery of trophic links (food chains) which describe the flow of energy/biomass between species is one of these problems. Networks of trophic links (food webs) are important for explaining ecosystem structure and dynamics [2]. However, relatively few ecosystems have been studied through detailed food webs because finding out the predation relationships between the many hundreds of species in an ecosystem is difficult and expensive. Hence, any technique which can automate the discovery of trophic links from ecological data is highly desirable. Similar problems of network construction have been tackled in other complex systems, such as metabolic networks (e.g. [8]). In this paper we demonstrate that Abductive ILP can generate plausible and testable food webs from ecological data. In this approach the abductive predicate ‘eats’ is entirely undefined before the start of the learning process. This contrasts with



**Fig. 1.** A farmland food web. Networks of trophic links (food web) are the primary routes for translocation of energy/biomass between species at different levels and are important for explaining ecosystem structure and dynamics.

previous applications of Abductive ILP where partial, non-empty, definitions exist and the gaps are filled by abduced hypotheses. In this paper we also explore a new approach for estimating probabilities for hypothetical ‘eats’ facts based on their frequency of occurrence when random permutations of the training data (and hence different seeds for defining the hypothesis space) are considered. We empirically evaluate the hypothetical trophic networks using leave-one-out cross-validation tests on the observable data. The results of cross-validation tests for the networks with and without probabilities are presented. The proposed trophic networks have been also examined by domain experts and the results of comparison with the literature are presented.

[PARAGRAPH TO EXPLAIN GROUND ABDUCTIVE HYPOTHESES IN THE CONTEXT OF ML, ILP AND RELATIONAL GRAPH LEARNING]

[TOADD: struct of the paper]

## 2 Ecological problem and data

Ecosystems are structured by flows of energy/biomass between primary producer plants (autotrophs) and consumers (heterotrophs), such as invertebrates, mammals and birds. Networks of trophic links (food web) are the primary routes for translocation of energy/biomass between species at different levels. For example, Figure 1 shows a food web in a farmland in which the food chains start from plants, then herbivores then predators which eat herbivores or other predators and so on.

There is much concern about anthropogenic and natural degradation of ecosystems worldwide, and the knock-on consequences for ecosystem functioning. Much research that seeks to identify mechanisms of ecosystem change and

devise methods for mitigating its effects is being hampered by an inability to construct food webs efficiently and effectively. Put simply, it takes considerable time and effort to establish trophic relationships between the many hundreds of species in an ecosystem - this means that only a relatively few systems have been studied to any depth making it difficult to produce general theories about ecosystem change. A method that yields plausible and testable food webs from already collected field data would be a major step forward in the identification of rules about how ecosystems react when perturbed.

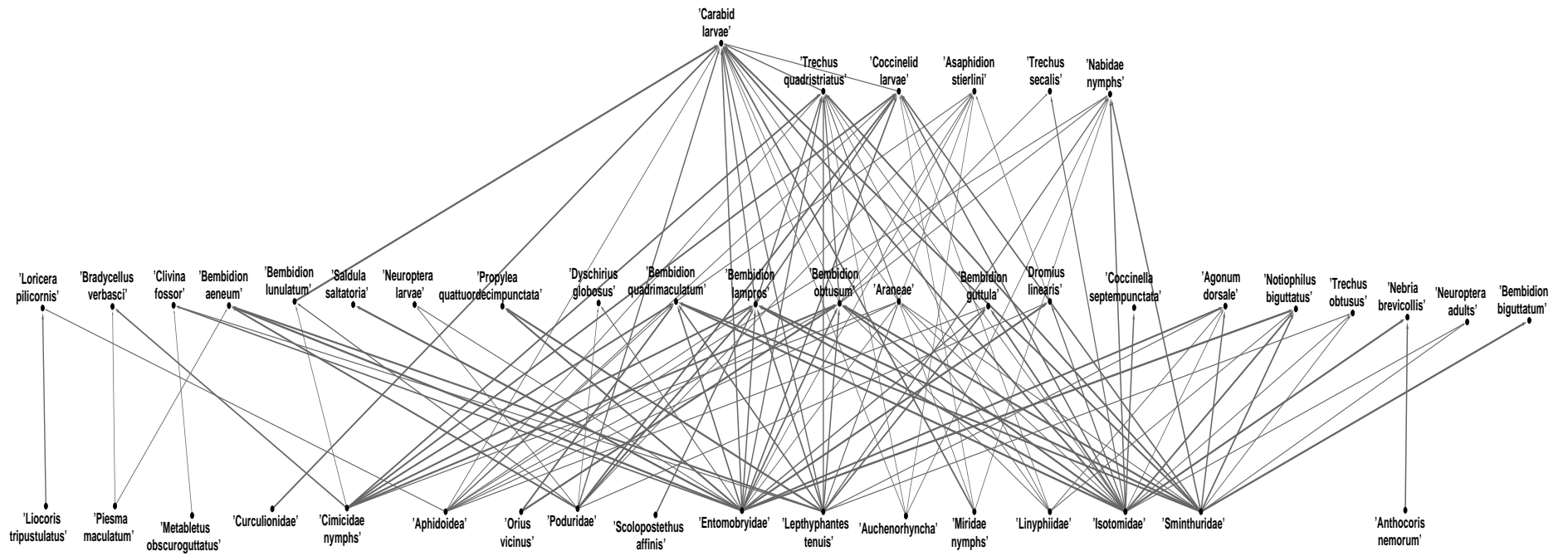
In this paper we try to answer the following question. Can machine learning be used to construct food webs from ecological data? We think the answer is yes and in fact the food web shown in Figure 1 has been learned from ecological data using Abductive ILP.

The training data we use comes from arable farmland where anthropogenic disturbance and farm management has led to great increases in crop productivity, but often at cost to biodiversity. Here, there is concern that the extent of biodiversity loss that has occurred (Benton, Vickery & Wilson 2003) might prevent ecosystem services, such as pollination and biological control, from being delivered (Loreau, Mouquet & Gonzalez 2003; Tscharnke et al. 2005). In this system, management disturbs trophic links, leading to the observed changes in diversity of the ecosystem (Caron-Lormier et al. 2009; 2011). The hope is that by evaluating trophic links, and their sensitivity to management, trophic networks might provide a mechanism for predicting ecosystem change (Macfadyen et al. 2009).

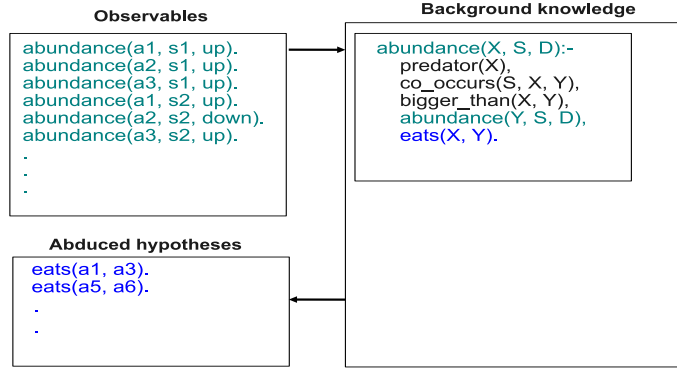
The data set was sampled from 257 fields across the UK in the Farm Scale Evaluations (FSE) of GM, herbicide tolerant (GMHT) crops. This national-scale experiment evaluated the change in weed plants and invertebrates between the current, conventional herbicide management of spring-sown Maize, Beet and Oilseed Rape and winter-sown Oilseed Rape, and the herbicide management of GMHT varieties of the same crops using a split-field design. We use data from the Vortis suction sampling protocol for epigeal invertebrates [6, 1] to calculate a treatment effect ratio. The counts from each conventional and GMHT half-field pair were converted to multiplicative treatment ratio,  $R$ , and as in [6, 1] treatment ratio values of  $R < 0.67$  and  $R > 1.5$  were regarded as important changes in count with direction of down (decreased) and up (increased), respectively. This information on up and down abundances is regarded as our primary observational data for the learning.

### 3 Machine learning of trophic links using Abductive ILP

The main role of abductive reasoning in machine learning and its use in the development of scientific theories [5] is to provide hypothetical explanations of the empirical observations. Then based on these explanations we try to inject back into the current scientific theory, new information that helps complete the theory. This process of generating abductive explanations and then updating in some way the theory with them can be repeated several times when new observa-



**Fig. 2.** Hypothetical trophic network (food web) constructed by A/ILP. Thickness of trophic links represent probabilities which are estimated based on the frequency of occurrence from 10 random permutations of the training data.



**Fig. 3.** Learning trophic links from ecological data using Abductive ILP.

tional data is made available. In many implementation of abductive reasoning, such as that of Prolog 5 [7], which is used in this paper, the approach taken is to choose an explanation that best generalises under some form of inductive reasoning (e.g. simplest explanation approximated by compressibility). We refer to this approach as Abductive ILP (A/ILP). We believe that ecological data in this study fulfil the conditions for the use of A/ILP: firstly, the given background knowledge is incomplete; and secondly, the problem requires learning in the circumstance in which the hypothesis language is disjoint from the observation language. In our problem, the set of observable data can be represented by predicate  $abundance(X, S, up)$  (or  $abundance(X, S, down)$ ) expressing the fact that the abundance of  $X$  at site  $S$  is *up* (or *down*). This information is compiled from FSE data as described in Section 2. The knowledge gap that we initially aim to fill is a predation relationship between species. Thus, we declare abducible predicate  $eats(X, Y)$  capturing the hypothesis that species  $X$  eats species  $Y$ . In order to use abduction, we also need to provide the rules which describe the observable predicate in terms of the abducible predicate. An example of such a rule is shown below.

```

abundance(X, S, up):-
  predator(X),
  co_occurs(S, X, Y),
  bigger_than(X, Y),
  abundance(Y, S, up),
  eats(X, Y).

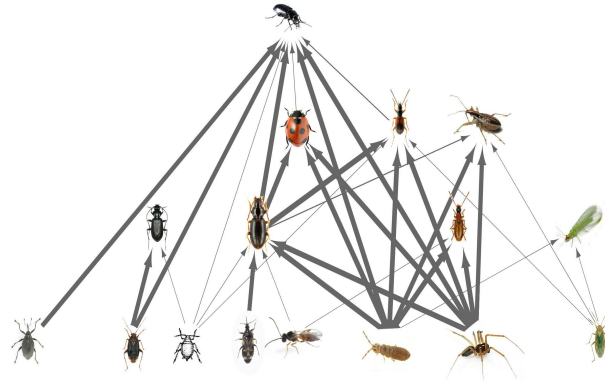
```

Similarly, a rule for  $abundance(X, S, down)$  can be defined. This Prolog rule expresses the inference that following a perturbation in the eco-system (caused by the management), the increased (or decreased) abundance of species  $X$  at site  $S$  can be explained by the fact that  $X$  eats species  $Y$  which is further down in

the food chain and the abundance of species  $Y$  is increased (or decreased). It also includes additional conditions to constraint the search for abducible predicate  $eats(X, Y)$ , i.e.  $X$  should be a predator,  $X$  and  $Y$  should co-occur and that  $X$  should be bigger than  $Y$ . Predicates  $predator(X)$ ,  $co\_occurs(S, X, Y)$  and  $bigger\_than(X, Y)$  are provided as part of the background knowledge. Given this model and the observable data, Progol 5 generates a set of ground abductive hypotheses in the form of ‘eats’ relations between species as shown in Figure ??.

This set of ground hypotheses can be visualised as a network of trophic links (food webs) as shown in Figure 2. In this network a ground fact  $eats(a, b)$  is represented by a trophic link from  $b$  to  $a$ .

[ADD TEXT FOR FUNCTIONAL GROUPS FOOD WEB]



**Fig. 4.** Food web between functional groups. Each node is represented by a species which can be viewed as an archetype for the functional group.

## 4 Hypothesis Frequency Estimation (HFE)

In order to get probability estimates for hypotheses, we use a technique which is based on direct sampling from the hypotheses space. In some ILP systems, including Progol, training examples also act as seeds to define the hypotheses space (e.g. a most specific clause is build from the next positive example). Hence, different permutations of the training examples define different parts of the hypothesis space. We use this property to sample from the hypothesis space by random permutations of the training data. Probability of ground hypotheses can be estimated based on the frequency of occurrence when random permutations of the training data (and hence different seeds for defining the hypothesis space) are considered. Using this technique, the thickness of trophic links in Figure 2 represent probabilities which are estimated based on the frequency of occurrence from 10 random permutations of the training data. This

probabilistic trophic network can be also represented using standard probabilistic representations in ILP such as Stochastic Logic Programs (SLPs) [?]slp) or ProbLog [?]problog). For this we can use relative frequencies in the same way probabilities are used in SLPs or ProbLog. We can then use the probabilistic inferences based on these representations to estimate probabilities. For example, the probability  $p(abundance(a, s, up))$  can be estimated by relative frequency of hypotheses which imply  $a$  at site  $s$  is  $up$ . Similarly,  $p(abundance(a, s, down))$  can be estimated and by comparing these probabilities we can decide to predict whether the abundance is up or down. We have used this method in leave-one-out experiments in Section 5 to measure the predictive accuracies of probabilistic trophic networks.

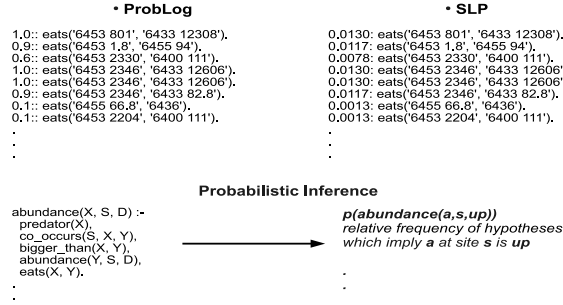


Fig. 5. Probabilistic representation and inference.

## 5 Empirical evaluation

The purpose of the experiments in this section is to empirically evaluate hypothetical trophic networks constructed from real ecological data using the methods described in the previous sections. In these experiments, firstly we determine if trophic networks can be learned from real ecological data such that their predictive accuracies on unseen data are higher than the default accuracy of the majority class. Secondly, we examine if using probabilities estimated by the permutation based HFE method (see Section 4) leads to increased predictive accuracy and also if our probability estimation method is comparable with a state-of-the-art probability estimation method from probabilistic ILP (i.e. LeP-robLog[]). Thirdly, we examine if we can improve the predictive accuracies by generalising over trophic interactions between individual species into trophic interactions between functional groups (see Section 3).

---

```

for  $t$  in [1..300] do
  for  $j=1$  to  $n$  do
     $Ts_{tj}$  = a randomly selected example from  $E_t$ 
     $Tr_{tj} = E_t - Ts_{tj}$ 
    for  $i$  in (25,50,75,100) do
       $Tr_{tji} = i$  training example randomly sampled from  $Tr_{tj}$ 
    end
  end
end
for  $t$  in [1..300] do
  for  $j=1$  to  $n$  do
    for  $i$  in (25,50,75,100) do
       $H_{tji}$  = learned hypotheses using the training set  $Tr_{tji}$ 
       $A_{tji}$  = predictive accuracy of  $H_{tji}$  on the test set  $Ts_{tj}$ 
    end
  end
end
for  $i$  in (25,50,75,100) do
  Plot average of  $A_{tji}$  versus  $i$  ( $j \in [1..n]$  and  $t \in [1..300]$ )

```

---

**Fig. 6.** Experimental method.  $E_t$  is the set of abundance data at site  $t$  and in this experiment  $n = 10$ . Variable  $i = (25, 50, 75, 100)$  represents the varying size of a randomly chosen training set in a leave-one-out test strategy. [INCOMPLETE]

## 5.1 Experiment 1

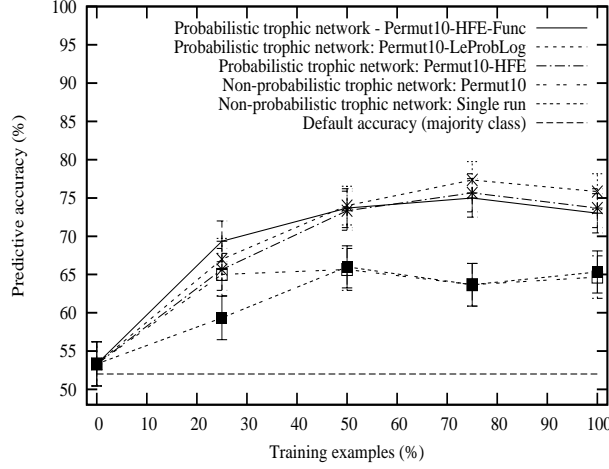
**Null hypothesis 1:** A trophic network with predictive accuracy more than the default accuracy cannot be constructed from ecological data using Abductive ILP.

**Materials and methods** In this experiment Progol 5.0<sup>4</sup> is used to generate abductive hypotheses (i.e. ‘eats’ relations between species) from observable data (i.e. up/down abundance of species at different sites). The observable data has been compiled from FSE data as described in Section 2. The up/down abundance of species at different sites are then encoded as predicates *abundance*( $X, S, up$ ) and *abundance*( $X, S, down$ ). The background knowledge includes information about sites and species and Prolog rules for *abundance* as well as predicates *predator*( $X$ ), *co\_occurs*( $S, X, Y$ ) and *bigger\_than*( $X, Y$ ) as described in Section 3. In order to empirically evaluate the hypothetical trophic networks, we use leave-one-out cross-validation test on the observable data for species in the network. This cross-validation test has been done by leaving out the abundance of each predator at each site and then trying to predict whether the abundance is up or down, given the trophic network generated from the rest of the data. This experimental method is detailed in Figure 6.

---

<sup>4</sup> Available from: <http://www.doc.ic.ac.uk/~shm/Software/progol5.0/>





**Fig. 7.** Predictive accuracies of probabilistic trophic network vs. non-probabilistic networks from leave-one-out cross-validation tests.

**Results and discussion** Figure 7 compares the predictive accuracy of non-probabilistic networks, i.e. networks generated from a single run or from 10 random permutations as well as probabilistic networks. In all cases the predictive accuracies are significantly higher than the default accuracy of the majority class (i.e. *down* for 53% of all abundance data). We can thus refute null hypothesis 1.

## 5.2 Experiment 2

**Null hypothesis 2:** Using the permutation based HFE method for estimating probabilities of trophic links does not lead to increased predictive accuracies compared to the non-probabilistic trophic network.

**Null hypothesis 3:** The predictive accuracy of a trophic network with probabilities estimated using the permutation based HFE method is significantly lower than the trophic network with probabilities estimated using LeProbLog.

**Materials and methods** As in Experiment 1, we use leave-one-out cross-validation test on the observable data for species in the network, i.e. leaving out the abundance of each predator at each site and trying to predict whether the abundance is *up* or *down*, given the trophic network generated from the rest of the data. For the trophic network with probabilities, we first need to calculate the relative frequencies of hypotheses which imply that the abundance of the test example is *up* or *down*. Let  $p(abundance(a, s, up))$  be the relative frequency of hypotheses which imply the abundance of  $a$  at site  $s$  is *up* and  $p(abundance(a, s, down))$  is defined analogously. If  $p(abundance(a, s, up)) > p(abundance(a, s, down))$  then we predict that the abundance of the test example

$a$  is *up* and otherwise it is *down*.  $p(abundance(a, s, up))$  and  $p(abundance(a, s, down))$  can be calculated using the probabilistic inference in SLPs or ProbLog given the rules for  $abundance(X, S, up)$  and  $abundance(X, S, down)$  and the relevant probabilistic representation of the trophic network in SLPs or ProbLog as described in Section 4.

**Results and discussion** As shown in Figure 7 the predictive accuracies for the non-probabilistic networks are significantly lower than the probabilistic networks when more than 50% of the training data are provided. Hence, we can reject null hypothesis 2. The predictive accuracies for the trophic network with probabilities estimated using LeProblog (Permut10-LeProbLog) are generally higher than the predictive accuracies of the trophic network with probabilities estimated using the permutation based HFE method (Permut10-HFE). In order to see if the differences are statistically significant we apply a t-test. The p-values from the t-test are 0.14, 0.28, 0.083 and 0.075 when 25%, 50%, 75% and 100% of training examples are used respectively. These p-values suggest that the null hypothesis 3 can also be rejected (significance level of 0.05).

### 5.3 Experiment 3

**Null hypothesis 4:** A trophic network between functional groups constructed by generalising trophic interactions over individual species does not have a higher predictive accuracy compared to the trophic network for individual species.

**Results and discussion** The predictive accuracies for the probabilistic network for functional groups (Permut10-HFE-Func) are shown in Figure 7. The difference between the predictive accuracies of the probabilistic network for the individual species (Permut10-HFE) and the network for functional groups (Permut10-HFE-Func) are not significant for 50%, 75% and 100% of training examples. The predictive accuracy of Permut10-HFE-Func for 25% is significantly higher than the predictive accuracy of Permut10-HFE (p-value of 0.004 from t-test). The null hypothesis 4 can therefore be refuted. This result suggest that when the number of training examples are limited, the network for functional groups has a higher predictive accuracy compared to network for the individual species.

## 6 Ecological evaluation

The trophic network in Figure 2 has been examined by the domain experts and corroboration of many of the links in the literature have been found. Table 2c is a tabular representation for some prey (columns) and predator (rows) species combination in Figure 2. Each pairwise hypothesised link has a strength (i.e. frequency between 1 to 10) followed by references (in square brackets) in the literature (listed in Appendix) supporting the link. In this table, only prey/predators

	Anthrenus nemorum	Bembidion lampros	Bembidion lunulatum	Bembidion obtusum	Cimicidae nymphs	Curculionidae	Entomobryidae	Isotomidae	Lepthyphantes tenuis	Licoris tripustulatus	Miridae nymphs	Orius vicinus	Poduridae	Scolopostethus affinis	Sminthuridae
Agonum dorsale							9 [13]	4	5 [6]						10 [13]
Bembidion aeneum							10 [11]		9 [6]				9 [11]		
Bembidion biguttatum															10 [11]
Bembidion guttula							7 [11]	10 [11]	9 [6]						
Bembidion lampros					9		10 [11]	10 [11]	10 [6]				10 [11]	9	10 [11]
Bembidion obtusum					9		10 [11]	10 [11]	10 [6]			10			10 [11]
Bembidion quadrimaculatum					9		10 [11]	10 [11]	9 [6]			10			9 [11]
Bradyobolus verbasci					8										
Carabid larvae		9	10	3		9	10 [2]	10 [2]	9		10		9 [2]		10 [2]
Citvina fossor							7 [12,5]		2						
Coccinellid larvae		9			9		10 [14,16]	9 [14,16]			1		10 [14,16]		9 [14,16]
Coccinella septempunctata								10 [14,16]							
Dromius linearis							10 [1]	7					3		
Loricera pilicornis										9					
Nabidae nymphs				3				10	7		2 [10,7]				10 [9]
Nebria brevicollis	10							10 [15]							
Notophilus biguttatus							10 [4,11,3]	10 [4,11,3]							10 [4,11,3]
Propylea quatuordecimpunctata							10 [14,16]		10 [16]						
Salidula saltatoria							10 [8]								
Trechus quadristriatus		9		9	9		9 [15,3]	4 [15,3]	9 [6]		2				10 [15,3]
Trechus secalis				2				4 [15,3]							

**Fig. 8.** Tabulated trophic links for some prey (columns) and predator (rows) species combination in Figure 2. Each pairwise hypothesised link has a strength (i.e. frequency between 1 to 10) followed by references (in square brackets) in the literature (listed in Appendix) supporting the link.

are shown which have at least one link with strength more than or equal to 7. This table shows that many of the links, suggested by the model, are corroborated by the literature. In particular, links in the model ascribed with high frequency are shown to correspond well with those having multiple references in the literature. Table 9 shows the correspondence between the frequencies of hypothetical trophic links, the number of multiple references and the total number of references in the literature. We use the numbers in this table to test the following null hypothesis.

**Null hypothesis 5:** The frequency of hypothetical trophic links in the learned trophic network are not correlated with the number of references in the literature for those trophic links.

Figure 9 shows the correlation between frequencies and the total number of references. If we use Spearman’s correlation between the frequencies and the total number of references then the  $\rho$  value and the  $p$ -value are 0.77 and 0.009 respectively. Hence, the hypothesis 5 is refuted.

According to Table 8 in some cases novel high frequency links are suggested, which could be tested. For example there are several links with frequency 10 but with no reference in the literature. New experimental studies are needed to test these potential novel hypotheses.

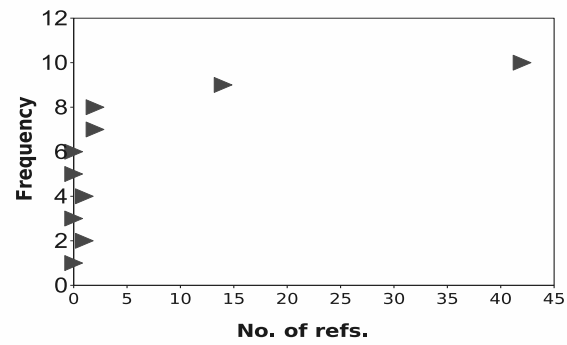
## 7 Discussion and related work

## 8 Conclusions

We find that machine learning, using A/ILP, produced a convincing food web from sample ecological data. Many of the important abduced trophic links are supported either by the literature or the expert knowledge of agricultural ecologists. This food web representing probabilistic interactions between species can readily be interpreted by Ecologists and the logical framework for learning trophic links can be openly discussed, a priori, and the hypothesised links are not an abstract, statistical product of the data. Two aspects of the use of A/ILP in this paper are particularly novel. Firstly, unlike previous applications of A/ILP, the abductive predicate ‘eats’ is entirely undefined before the start of the learning process. The second novel aspect of the approach relates to the assignment of probabilities to hypothetical ‘eats’ facts based on their frequency of occurrence when randomly sampling the hypothesis space. The resulting probabilistic network is a compact summary of the hypothesis space with a posterior distribution which could be viewed as a Bayes predictor. The results of cross-validation tests suggest that the trophic networks with probabilities have higher predictive accuracies compared to the networks without probabilities. In this paper we have reported the predictive accuracies for binary classification. However, we have also used expected utilities implemented as Decision-Theoretic Logic Programs (DTLPs) [3] for estimating R values (treatment effect ratio as

Freq.	1 Ref.	2 Ref.	3 Ref.	Total refs.
1	0	0	0	0
2	0	1	0	1
3	0	0	0	0
4	0	1	0	1
5	0	0	0	0
6	0	0	0	0
7	1	1	0	2
8	0	1	0	2
9	8	3	0	14
10	23	5	3	42

(a)



(b)

**Fig. 9.** a) Correspondence between the frequencies of hypothetical trophic links, the number of multiple references and the total number of references in the literature b) The correlation between frequencies and the total number of references. Spearman's correlation  $\rho$  value is 0.77 with  $p$ -value 0.009.

described in Section 2). Initial results suggest that using probabilities leads to reduced mean square errors when estimating R values in cross-validation tests. The probabilistic trophic network together with the expected utility approach can be viewed as a Decision-Theoretic representation which we call an Acyclic Expectation Network (AEN). We intend to study different aspects of this representation in a follow up paper.

[EXTEND CONCLUSIONS WITH RESULTS ON FUNC. GROUPS] This is to our knowledge the first time that machine learning and in particular ILP has been used for the construction of food webs from ecological data. We believe that this paper represents a major breakthrough for ecosystem and food web research that will lead to important advances in ecological theory and improved management of ecosystems under environmental change. We envisage immediate opportunities to apply this knowledge to optimize the delivery of food and other ecosystem services from agricultural land.

## Acknowledgements

We would like to thank S. Dunbar for his encouragement and support. We also thank G. Caron-Lormier, J. Chen, V. Lesk for helpful discussions and the anonymous referees for useful comments. This work is funded by Syngenta Ltd.

## References

1. D.A. Bohan, C.W.H. Boffey, D.R. Brooks, S.J. Clark, A.M. Dewar, L.G. Firbank, A.J. Haughton, C. Hawes, M.S. Heard, M.J. May, et al. Effects on weed and invertebrate abundance and diversity of herbicide management in genetically modified herbicide-tolerant winter-sown oilseed rape. *Proceedings of the Royal Society B: Biological Sciences*, 272(1562):463, 2005.
2. G. Caron-Lormier, D.A. Bohan, C. Hawes, A. Raybould, A.J. Haughton, and R.W. Humphry. How might we model an ecosystem? *Ecological Modelling*, 220(17):1935–1949, 2009.
3. J. Chen and S. Muggleton. Decision-theoretic logic programs. In *Proceedings of ILP*, 2009.
4. T.G. Dietterich. Machine learning in ecosystem informatics and sustainability. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence. Pasadena, Calif.: IJCAI*, pages 8–13, 2009.
5. P.A. Flach and A.C. Kakas. *Abduction and Induction: Essays on their relation and integration*. Springer Netherlands, 2000.
6. A.J. Haughton, G.T. Champion, C. Hawes, M.S. Heard, D.R. Brooks, D.A. Bohan, S.J. Clark, A.M. Dewar, L.G. Firbank, J.L. Osborne, et al. Invertebrate responses to the management of genetically modified herbicide-tolerant and conventional spring crops. ii. within-field epigeal and aerial arthropods. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1439):1863, 2003.
7. S. Muggleton and C. Bryant. Theory completion using inverse entailment. *Inductive Logic Programming*, pages 130–146, 2000.
8. A. Tamaddoni-Nezhad, R. Chaleil, A. Kakas, and S. Muggleton. Application of abductive ILP to learning metabolic network inhibition from temporal data. *Machine Learning*, 64(1):209–230, 2006.

## Appendix: References used for corroboration in Table 2c

1. K.N.A. Alexander. The invertebrates of living and decaying timber in britain and ireland—a provisional annotated checklist. *English Nature Research Reports*, 467:1–142, 2002.
2. T. Bauer. Prey-capture in a ground-beetle larva. *Animal Behaviour*, 30(1):203–208, 1982.
3. J.R. Bell, R. Andrew King, D.A. Bohan, and W.O.C. Symondson. Spatial co-occurrence networks predict the feeding histories of polyphagous arthropod predators at field scales. *Ecography*, 33(1):64–72, 2010.
4. K. Berg. *The role of detrital subsidies for biological control by generalist predators evaluated by molecular gut content analysis*. PhD thesis, Universitäts- und Landesbibliothek Darmstadt, 2007.

5. K. Desender and M. Pollet. Ecological data on *clivina fossor* (coleoptera, carabidae) from a pasture ecosystem. ii. reproduction, biometry, biomass, wing polymorphism and feeding ecology. *REV. ECOL. BIOL. SOL.*, 22(2):233–246, 1985.
6. A. Dinter. Intraguild predation between erigonid spiders, lacewing larvae and carabids. *Journal of Applied Entomology*, 122(1-5):163–167, 1998.
7. J.D. Lattin. Bionomics of the nabidae. *Annual review of entomology*, 34(1):383–400, 1989.
8. J. Lindsey. Ecology of *Commanster*, <http://www.commanster.eu/commanster/insects/bugs/spbugs/saldula.saltatoria.html>.
9. X. Pons, B. Lumbierres, and R. Albajes. Heteropterans as aphid predators in inter-mountain alfalfa. *European Journal of Entomology*, 106(3):369–378, 2009.
10. C.W. Schaefer and A.R. Panizzi. *Heteroptera of economic importance*. CRC, 2000.
11. K.D. Sunderland. The diet of some predatory arthropods in cereal crops. *Journal of Applied Ecology*, 12(2):507–515, 1975.
12. K.D. Sunderland, G.L. Lovei, and J. Fenlon. Diets and reproductive phenologies of the introduced ground beetles *harpalus-affinis* and *clivina-australasiae* (coleoptera, carabidae) in new-zealand. *Australian Journal of Zoology*, 43(1):39–50, 1995.
13. S. Toft. The quality of aphids as food for generalist predators: implications for natural control of aphids. *European Journal of Entomology*, 102(3):371, 2005.
14. B.D. Turner. Predation pressure on the arboreal epiphytic herbivores of larch trees in southern england. *Ecological Entomology*, 9(1):91–100, 1984.
15. D.J. Warner, L.J. Allen-Williams, S. Warrington, A.W. Ferguson, and I.H. Williams. Mapping, characterisation, and comparison of the spatio-temporal distributions of cabbage stem flea beetle (*psylliodes chrysocephala*), carabids, and collembola in a crop of winter oilseed rape (*brassica napus*). *Entomologia experimentalis et applicata*, 109(3):225–234, 2003.
16. D.C. Weber and J.G. Lundgren. Assessing the trophic ecology of the coccinellidae: their roles as predators and as prey. *Biological Control*, 51(2):199–214, 2009.