InfoVeg - Bug #2639

Inferred relationships too broad

11/01/2006 07:36 AM - Robert Peet

Status:	Closed	**Start date:**	11/01/2006
Priority:	Immediate	**Due date:**	
Assignee:	Xianhua Liu	**% Done:**	0%
Category:	atlas	**Estimated time:**	0.00 hour
Target version:	Unspecified	**Spent time:**	0.00 hour
Bugzilla-Id:	2639		

Description

The Carya glabra example reveals a new bug. All of USDA and RAB map as ambiguous. This should not be the case because the concepts of C glabra invoked by RAB, USDA, K, & Weakley are all the same. The crux of the problem is that the concept relationships shows
Carya glabra (Alan Weakley) <= Carya glabra (RAB)
Carya glabra (Alan Weakley) <= Carya glabra (K)

The root of the problem is that Alan did not map concepts for full species when there were vars. Xianhua had to create a script that would deduce the relationships for the higher order taxa. His script was not able to decide whether the three vars of Carya glabra recognized summed to the full species in RAB and K, or was a subset. It is not immediately clear to me how to do this, but perhaps because the RAB glabra did not map onto any other Weakley concepts, we could conclude that the relationship was = and not <. I propose Xianhua tries to modify his script to correct the concept relationships in this fashion.

History

#1 - 11/05/2006 09:04 PM - Robert Peet

In the case of a species with varieties in Weakley, we must deduce the relationship of the Weakley species-level taxon (both to specific concepts and to nominal concepts) as Weakley only mapped the vars. The current algorithm takes a conservation approach and states that the full Weakley species is <= the full species in other sources and as the nominal, even when the related concepts and the nominal concept show no relationships to other Weakley concepts. Examples are Quercus macrocarpa, Q. rubra, and Aesculus glabra. In each case the relationship deduced by our algorithm is Aus beus sec Weakley <= Aus beus sec other, and <= Aus beus nominal. We should check to see if non-Weakley concepts point in addition to other Weakley concepts, and if not assume the above relationship is = rather than <=. Elsewise we get unambiguous specimens mapped as ambiguous owing to a nominal with a <= relationship to the Weakley concept.

#2 - 11/12/2006 01:11 PM - Xianhua Liu

Rule used here is (rule 24 in the rule document):

IF all children of A < B THEN A<=B

Original relationships:

Carya glabra var. glabra (Alan Weakley) < Carya glabra (RAB)
Carya glabra var. hirsuta (Alan Weakley)< Carya glabra (RAB)
Carya glabra var. megacarpa (Alan Weakley)< Carya glabra(RAB)

Following the rule, we have: Carya glabra(Alan Weakley) <= Carya glabra(RAB).

I can change the rule to :

IF all children of A < B AND No relationships exit between any other concepts and B, THEN A=B.

For sure, this new rule will result in: Carya glabra(Alan Weakley) = Carya glabra(RAB), since there is no other relationships from Weakley's concepts to Carya glabra(RAB). But I am concerning that if we will risk introducing new problems by applying this new role where it is not in a case same to "Carya glabra". Do we have any exception?

#3 - 12/13/2006 01:45 PM - Robert Peet

I cannot immediately think of any exceptions. Let's immediately implement the new rule.

#4 - 12/16/2006 10:43 AM - xianhua liu

Fixed.

1. changed the inferring rule from:

"IF all children of A < B THEN A<=B" to 2 new rules:

I. "IF all children of A < B AND there are other concepts related to B THEN A<=B"
II. "IF all children of A < B AND no other concepts related to B THEN A=B"

2. Re-ran the algorithm with the new rules

3. updated the relationships at the herbarium server

#5 - 12/16/2006 10:43 AM - xianhua liu

fixed

#6 - 02/04/2007 08:44 PM - Robert Peet

We are making progress, but new problems keep appearing. Let's continue with the Carya glabra example.

There is an inferred rule that C.g. (weakley) <= C. g (FNA)
This should be C.g. (Weakley) < C.g. (FNA)
This is because we know that C ovalis (Weakley) < C.g. (FNA)
The inferring rules may need to be changed to cover these sorts of cases.

Another problem is that if you look at Quercus rubra you find
Q. rubra (Weakley) >= Q. rubra (nominal).
Nominal here is defined as the broadest possible usage, so we do not allow Q. rubra (W) to ever be > Q. rubra nominal.

[As an aside, another and more difficult problem is the inferred relationship
C.g. (Weakley) < C. ovalis (nominal)
The correct relationship is |
The root of the problem is that there is set of specimens that variously get called C.g. var hirsuta and C.ovalis var hirsuta. This does not cause a mapping issue, so I suggest we skip over it unless you see an obvious solution.]

#7 - 02/06/2007 05:21 PM - xianhua liu

[As an aside, another and more difficult problem is the inferred relationship
C.g. (Weakley) < C. ovalis (nominal)
The correct relationship is |
The root of the problem is that there is set of specimens that variously get
called C.g. var hirsuta and C.ovalis var hirsuta. This does not cause a
mapping issue, so I suggest we skip over it unless you see an obvious
solution.]

If the nominal concept is defined as the broadest possible usage, "<" might be a better choice.

#8 - 02/06/2007 06:04 PM - xianhua liu

Fixed

1. changed the inferring rule to allow C.g. (Weakley) < C.g. (FNA) and other relationships under the similar situation. The result can be checked now.

2. Changed the inferring rule for relationships to nominal concepts, which allows Q. rubra (W) = Q. rubra nominal and other relationships under the similar situation.

3. Java class to do the nominal inferring function has been changed and email sent to Hinar notifying the change. This change will not take effect until the class is updated on the herbarium server.

#9 - 09/08/2007 12:06 PM - Robert Peet

fixed

#10 - 03/27/2013 02:20 PM - Redmine Admin

Original Bugzilla ID was 2639