

## ΣΧΕΔΙΑΣΗ ΟΝΤΟΛΟΓΙΩΝ

(από εργασία του φοιτητή Παναγιωτόπουλου

Παναγιώτη στο άρθρο:

Ontology Development101: A Guide to Creating Your  
First Ontology:

[http://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](http://protege.stanford.edu/publications/ontology_development/ontology101.pdf) )

# Οντολογίες

- Υπάρχουν διάφοροι ορισμοί. Η οντολογία (ontology), ως μια τυποποιημένη περιγραφή ενός συγκεκριμένου τομέα γνώσης η οποία πρέπει να είναι αποδεκτή από μια ομάδα ατόμων, για να έχει νόημα η ύπαρξή της, έρχεται να καλύψει το πρόβλημα της πολυσημίας (πολλές λέξεις περιγράφουν το ίδιο αντικείμενο) και της αμφισημίας (μια λέξη περιγράφει διαφορετικά αντικείμενα). Μια οντολογία περιγράφει με ακρίβεια τις έννοιες που αποφασίζει η ομάδα να περιγράψει.
- Το πιο δύσκολο μέρος της κατασκευής της είναι οι «παραδοχές» που πρέπει να λάβουν υπόψη οι κατασκευαστές της για το περιεχόμενο της οντολογίας (ποιες έννοιες θα περιγραφούν και ποιες όχι, αφού δεν μπορούν να περιγραφούν τα πάντα), το σκοπό δημιουργίας της (ποιοι θα είναι οι χρήστες της και που θα τη χρησιμοποιήσουν), το είδος της οντολογίας που θα αναπτυχθεί (μια οντολογία που θα περιγράφει γενικούς όρους, όπως ο χρόνος και ο χώρος κ.τ.λ).
- Οι οντολογίες ως εννοιολογικές μορφοποιήσεις (conceptualizations) αποτελούν προϊόντα υποκειμενικής κρίσης, οπότε το ίδιο πεδίο ενδιαφέροντος είναι δυνατόν να περιγραφεί με διαφορετικούς τρόπους. Άρα οι οντολογίες που αναπτύσσονται αποτελούν ετερογενείς πηγές γνώσης.

# Οντολογίες–Λόγοι δημιουργίας

- Κοινή χρήση της δομής της πληροφορίας ανάμεσα σε ανθρώπους και μηχανές: αποτελεί και τον πιο διαδεδομένο λόγο ανάπτυξης οντολογιών. Όταν με τη βοήθεια μιας οντολογίας αναπτύσσονται διαφορετικές σελίδες που βασίζονται σε αυτήν, τότε οι εφαρμογές μπορούν να συνδυάσουν πληροφορίες για να απαντήσουν στις αναζητήσεις του χρήστη.
- Επαναχρησιμοποίηση της γνώσης μίας περιοχής: μία καλά ορισμένη οντολογία από μια ομάδα, μπορεί να επαναχρησιμοποιηθεί σε άλλες περιοχές και από άλλες ομάδες ή η ανάπτυξη μιας νέας οντολογίας μπορεί να βασιστεί στη συνένωση ήδη υπάρχοντων επιμέρους οντολογιών.
- Σαφής ορισμός των αξιωμάτων μιας γνωστικής περιοχής: σκοπός είναι να γίνονται ευκολότερα οι αλλαγές στην οντολογία εξαιτίας αλλαγών που προκύπτουν στο πεδίο της γνώσης. Από την άλλη μεριά ο σαφής ορισμός των εννοιών βοηθά νέους χρήστες να κατανοήσουν την υπάρχουσα γνώση.

# Οντολογίες–Λόγοι δημιουργίας

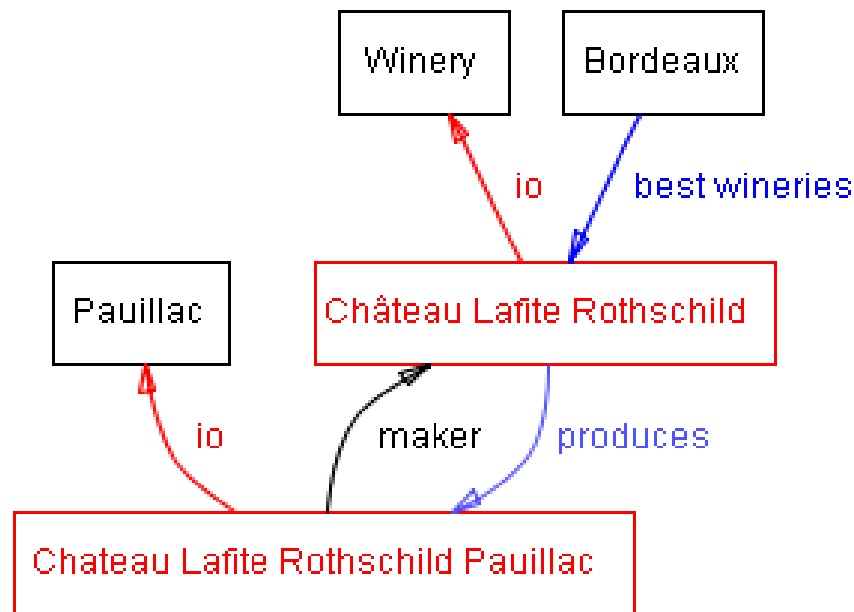
- Διαχωρισμός της γνώσης μιας περιοχής από τη λειτουργική γνώση: μια διαδικασία η οποία περιγράφεται για κάποιες συσκευές, μπορεί να χρησιμοποιηθεί αυτούσια για κάποιες άλλες, αν αυτές προστεθούν στην χρησιμοποιούμενη οντολογία, χωρίς όμως να χρειάζονται αλλαγές στον ορισμό της διαδικασίας.
- Ανάλυση της γνώσης: η ανάλυση της γνώσης καθίσταται δυνατή με χρήση όρων που είναι διαθέσιμοι και οι οποίοι πρέπει να έχουν παραχθεί από μια τυπική διαδικασία ανάλυσης.

# Δομικά στοιχεία οντολογίας στο protege

- Κλάσεις (classes). Οι κλάσεις αναπαριστώνται με τις έννοιες (concepts). Μια έννοια μπορεί να είναι οτιδήποτε για το οποίο μπορεί να ειπωθεί κάτι και μπορεί να είναι η περιγραφή μιας εργασίας, μιας λειτουργίας, μιας ενέργειας, μιας ιδέας, μιας κρίσης, κλπ.
- Slots: Περιγραφή διαφόρων χαρακτηριστικών και ιδιοτήτων των κλάσεων.
- Facets: Περιορισμοί που μπαίνουν στα slots.
- Στιγμιότυπα (instances): Τα στιγμιότυπα αναπαριστούν συγκεκριμένα στοιχεία. Για παράδειγμα, ο φοιτητής με το όνομα «Παναγιώτης» είναι στιγμιότυπο της κλάσης «φοιτητής».
- Σχέσεις (relations). Οι σχέσεις εκφράζουν ένα είδος αλληλεπίδρασης μεταξύ των εννοιών ενός πεδίου ενδιαφέροντος, για παράδειγμα subclass-of, is-a, κλπ.
- Συναρτήσεις (functions). Οι συναρτήσεις εκπροσωπούν μια ειδική περίπτωση σχέσης. Για παράδειγμα, η τιμή-μεταχειρισμένου-αυτοκινήτου μπορεί να προσδιορίζεται ως συνάρτηση της αρχικής τιμής του αντίστοιχου καινούργιου αυτοκινήτου, του μοντέλου και των χαρακτηριστικών του αυτοκινήτου, των χιλιομέτρων που έχει διανύσει, κλπ.

# Δημιουργώντας μια οντολογία...

- Δημιουργώντας μια οντολογία χρειάζεται να ορίσουμε τις κλάσεις μέσα σε αυτή.
- Ταξινόμηση των κλάσεων αυτών (υπο-κλάσεις, υπερ-κλάσεις)
- Καθορισμός των ιδιοτήτων (slots) και των τιμών που μπορούν να πάρουν.



Στο σχήμα φαίνονται κάποιες κλάσεις (σε μαύρο πλαίσιο), στιγμιότυπα (σε κόκκινο πλαίσιο) και σχέσεις μεταξύ τους στο πεδίο των κρασιών (wine). Οι απ'ευθείας σύνδεσμοι αναπαριστούν ιδιότητες και εσωτερικούς συνδέσμους, όπως "instance-of" και "subclass-of".

# Βήματα για τη δημιουργία μιας οντολογίας

- Βήμα 1:Καθορισμός πεδίου και σκοπός της οντολογίας
  - Ποιο είναι το πεδίο που θα καλύψει η οντολογία;
  - Για ποιο λόγο πρόκειται να χρησιμοποιήσουμε την οντολογία;
  - Για τι είδους ερωτήσεις η οντολογία θα παρέχει απαντήσεις;
  - Ποιός θα χρησιμοποιήσει και θα διατηρήσει την οντολογία αυτή;
- Βήμα 2:Δυνατότητα επαναχρησιμοποίησης υπάρχουσας οντολογίας
  - Για παράδειγμα αν θέλουμε να δημιουργήσουμε την οντολογία wines και κάνουμε import την οντολογία French wines, τότε όχι μόνο έχουμε μια κατηγοριοποίηση των γαλλικών κρασιών αλλά και μια πρώτη ιδέα για τα χαρακτηριστικά των κρασιών τα οποία τα διαφοροποιούν το ένα απο το άλλο.

# Βήματα για τη δημιουργία μιας οντολογίας

- Βήμα 3: Επιλογή των στοιχείων μιας οντολογίας.

Για παράδειγμα δημιουργώντας μια οντολογία wine, χρειάζεται να ορίσουμε κάποια χαρακτηριστικά τα οποία περιγράφουν την οντολογία wine, όπως το grape, location, colour, food accompaniment

Σημείωση: Αρχικά μπορούμε να δημιουργήσουμε μια λίστα απο όρους, χωρίς να σκεφτόμαστε ποια απ'τα αντικείμενα αλληλοκαλύπτονται (κοινή θεματολογία) ή αν αυτά αποτελούν κλάση ή ιδιότητα (slot).



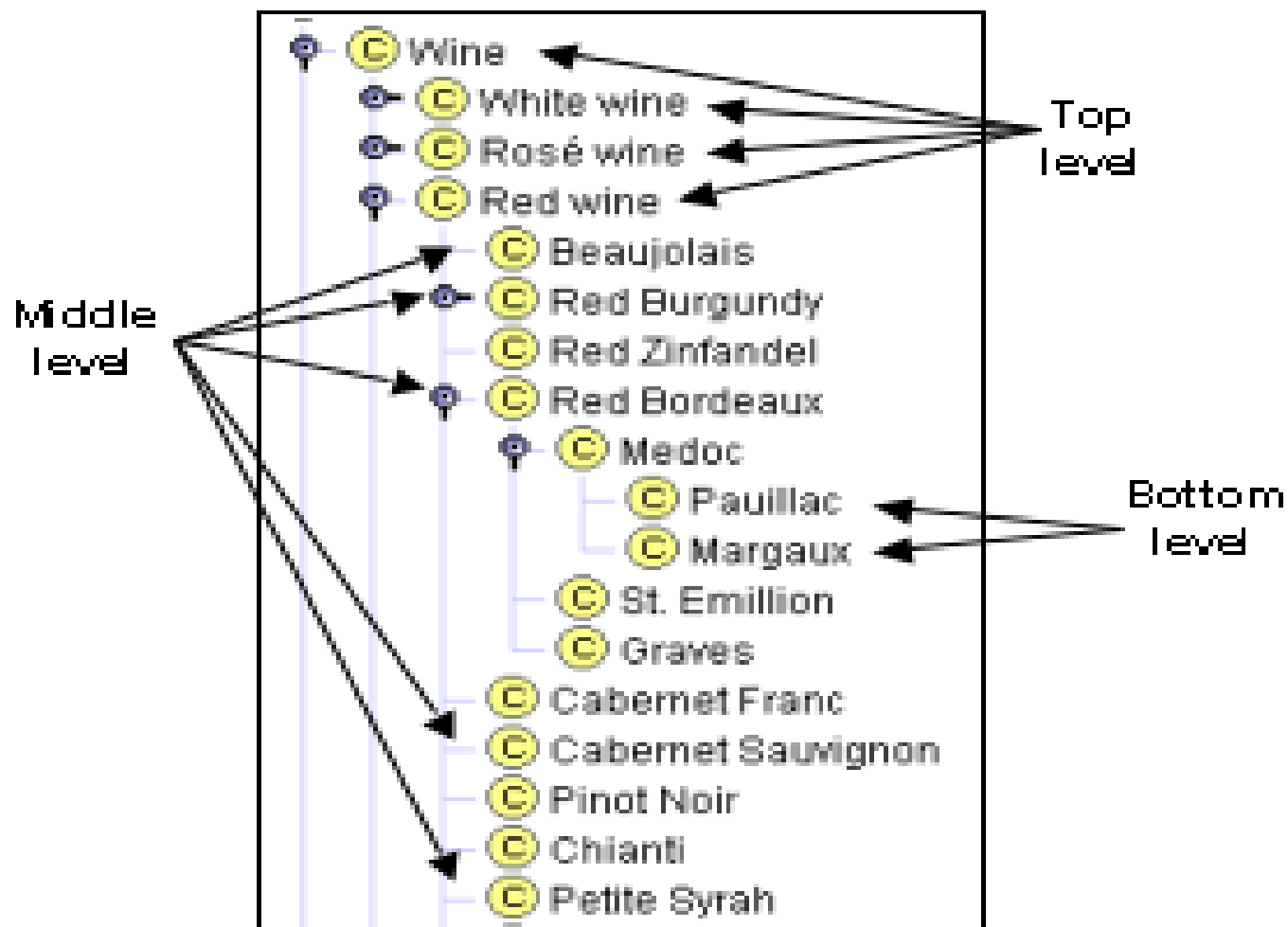
# Βήματα για τη δημιουργία μιας οντολογίας

- Βήμα 4:Καθορισμός των κλάσεων και της ιεραρχίας κλάσεων

- Στην ονομαζόμενη top-down ανάπτυξη της οντολογίας εισάγουμε την κλάση που θα περιγράψουμε και κάποιες βασικές υποκλάσεις. Για παράδειγμα ορίζουμε την κλάση wine και έπειτα τις υποκλάσεις white wine, red wine, rose wine, ενώ μια περαιτέρω κατηγοριοποίηση της κλάσης red wine σε Red Burgundy, Cabernet Sauvignon κλπ.
- Στην ονομαζόμενη bottom-up ανάπτυξη ορίζουμε τις κλάσεις του τελευταίου επιπέδου (μπορούμε να το παρομοιάσουμε με τα φύλλα ενός δέντρου), το οποίο αποτελεί υποκλάση μιας πιο γενικής κλάσης. Για παράδειγμα, αρχικά δημιουργούμε τις κλάσεις Pauillac και Margeux. Έπειτα δημιουργούμε μια υπερκλάση Medoc για τις δύο αυτές κλάσεις, η οποία είναι υποκλάση της κλάσης Red Bordeaux.
- Τέλος, μπορεί να γίνει συνδυασμός των δύο παραπάνω. Κάποιος μπορεί να ξεκινήσει απ'την top-level ανάπτυξη(Wine), έπειτα να χρησιμοποιήσει bottom-level(Margeux) και να την συνδέσει με την middle-level ανάπτυξη(Medoc)

# Βήματα για τη δημιουργία μιας οντολογίας

- ▶ Παράδειγμα οντολογίας στο protege



# Βήματα για τη δημιουργία μιας οντολογίας

- Η χρήση των παραπάνω μεθόδων βασίζεται σε υποκειμενικά κριτήρια. Για παράδειγμα αν κάποιος ενδιαφέρεται για την διαφοροποίηση των κρασιών σε είδη (λευκό, κόκκινο, ροζέ), τότε θα είναι προτιμότερο να χρησιμοποιήσει την top-down ανάπτυξη. Αν όμως ενδιαφέρεται για ποιο συγκεκριμένα στοιχεία (τίτλο κρασιού), τότε είναι προτιμότερο να χρησιμοποιήσει την bottom-level ανάπτυξη.
- Όποια ανάπτυξη και να επιλεγεί πρέπει να γίνει καθορισμός εκείνων των κλάσεων οι οποίες είναι ανεξάρτητες από άλλες, προσπαθώντας έτσι να γίνει ιεραρχική ταξινόμηση ελέγχοντας αν μια κλάση αποτελεί υποκλάση μιας άλλης. Π.χ το κρασί Pinot Noir είναι κόκκινο κρασί. Άρα θα αποτελεί απαραίτητα υποκλάση της κλάσης Red Wine.

# Βήματα για τη δημιουργία μιας οντολογίας

- ▶ Τα slots της κλάσης wine και τα facets των slots, όπως εμφανίζονται στο protege

Template Slots				V	V	C	X	+	-
Name	Type	Cardinality	Other Facets						
S body	Symbol	single	allowed-values={FULL,MEDIUM,LIGHT}						
S color	Symbol	single	allowed-values={RED,ROSÉ,WHITE}						
S flavor	Symbol	single	allowed-values={DELICATE,MODERATE,STRONG}						
S grape	Instance	multiple	classes={Wine grape}						
S maker <sup>1</sup>	Instance	single	classes={Winery}						
S name	String	single							
S sugar	Symbol	single	allowed-values={DRY,SWEET,OFF-DRY}						

# Βήματα για τη δημιουργία μιας οντολογίας

- ▶ Βήμα 5: Καθορισμός των τιμών των κλάσεων-slots
- ▶ Οι κλάσεις από μόνες τους δεν μπορούν να δώσουν αρκετές πληροφορίες. Αυτό γίνεται με τον ορισμό των ιδιοτήτων τους (flavor, colour, sugar).
- Σε μια οντολογία slots μπορούν να αποτελούν: "εγγενείς" ιδιότητες(flavor), "εξωγενείς" ιδιότητες (name, area), σχέσεις μεταξύ ανεξάρτητων μελών της κλάσης και άλλων αντικειμένων (π.χ maker της Wine αναπαριστά την σχέση μεταξύ wine και winery).

# Βήματα για τη δημιουργία μιας οντολογίας

- Βήμα 6: Ορισμός όψεων (facets) των slots
- ▶ Τα slots μπορούν να έχουν όψεις (facets) όπως τύπο μεταβλητής, αριθμό μεταβλητών κτλ.
- ▶ –Slot cardinality: Ορίζει πόσες τιμές μπορεί να έχει ένα slot. π.χ το body της κλάσης Wine έχει single cardinality slot αφού ένα κρασί έχει ένα τύπο μπουκαλιού.
- ▶ –Slot-value type: π.χ String, Number, Boolean, Enumerated (για την γεύση μπορεί να είναι strong, moderate, delicate)
- ▶ –Slot τύπου στιγμιοτύπου(instance): Για παράδειγμα το slot produces της κλάσης Winery μπορεί να έχει στιγμιότυπα της κλάσης Wine.

# Βήματα για τη δημιουργία μιας οντολογίας

- Το παρακάτω σχήμα δείχνει την σχέση του slot produces με την κλάση Winery.

The screenshot shows a window titled "S produces" with the following fields and controls:

- Name:** A text box containing "produces".
- Documentation:** A text area containing "This slot contains the wines produced by a particular winery".
- Value Type:** A dropdown menu set to "Instance".
- Allowed Classes:** A list box containing "Wine" with a yellow circle icon to its left. There are "+" and "-" buttons to the right of the list.
- Cardinality:** Two rows of controls:
  - Row 1:  required, followed by "at least" and an empty text box.
  - Row 2:  multiple, followed by "at most" and an empty text box.

# Βήματα για τη δημιουργία μιας οντολογίας

- ▶ Domain και range ενός slot
- ▶ –Οι επιτρεπόμενες κλάσεις για slots τύπου instance ονομάζονται range του slot.
- ▶ –Οι κλάσεις στις οποίες επισυνάπτεται ένα slot ονομάζονται domain του slot (στο προηγούμενο παράδειγμα η κλάση Winery αποτελεί το domain του slot produces).
  
- Για να ορίσουμε το domain και το range ενός slot, βρίσκουμε τις πιο γενικές κλάσεις οι οποίες μπορούν να είναι domain ή range ενός slot.
- Αν υπάρχει μια λίστα κλάσεων οι οποίες ορίζουν το domain ή το range ενός slot, περιέχουν και υποκλάσεις τους, τότε αφαιρούμε τις υποκλάσεις αυτές.





# Βήματα για τη δημιουργία μιας οντολογίας

- Βήμα 7: Δημιουργία στιγμιοτύπων.
- ▶ -Για την δημιουργία ενός στιγμιοτύπου μιας κλάσης:  
(1)επιλέγουμε την κλάση, (2) δημιουργούμε ένα ανεξάρτητο στιγμιότυπο της κλάσης, (3) συμπληρώνουμε τις τιμές των slots.

The screenshot shows a window titled "Chateau Morgon Beaujolais (Beaujolais)". The interface is divided into several sections for defining the wine's properties:

- Name:** Chateau Morgon Beaujolais
- Area:** Beaujolais region
- Body:** LIGHT
- Color:** RED
- Maker:** Chateau Morgon
- Flavor:** DELICATE
- Sugar:** DRY
- Grape:** Gamay grape
- Tannin Level:** LOW

Each section has a dropdown menu and a "V" button, with some sections also having "+" and "-" buttons for adjustment.

# Ορισμός κλάσεων και ιεραρχία κλάσεων

- Επιβεβαίωση ότι η ιεραρχία των κλάσεων είναι σωστή
- ▶ -Η σχέση “is-a”: Η κλάση A είναι υποκλάση της B, αν οποιοδήποτε στιγμιότυπο της A είναι και στιγμιότυπο της B.
- ▶ -Μεταβατικότητα στις σχέσεις των κλάσεων: Αν η κλάση B είναι υποκλάση της A και η Γ είναι υποκλάση της B, τότε και η Γ είναι υποκλάση της B.
- ▶ -Αποφυγή κύκλων στις κλάσεις: Αυτό σημαίνει η κλάση A να έχει μια υποκλάση B και ταυτόχρονα η B να είναι υπερκλάση της A.
- ▶ -Η χρήση συνώνυμων για την δήλωση του ίδιου θέματος δεν αναπαριστά διαφορετικές κλάσεις.

# Ορισμός κλάσεων και ιεραρχία κλάσεων

Ανάλυση siblings στην ιεραρχία κλάσεων

- ▶ -Siblings είναι κλάσεις, οι οποίες είναι υποκλάσεις των ίδιων κλάσεων και πρέπει να βρίσκονται στο ίδιο επίπεδο γενικότητας
- ▶ -Εάν μια κλάση έχει ένα μόνο sibling, τότε ίσως είναι δεν είναι πλήρης.
- ▶ -Απ'την άλλη, εάν έχει αρκετά siblings, τότε ίσως χρειαστούν ενδιάμεσες κατηγορίες.



# Ορισμός κλάσεων και ιεραρχία κλάσεων

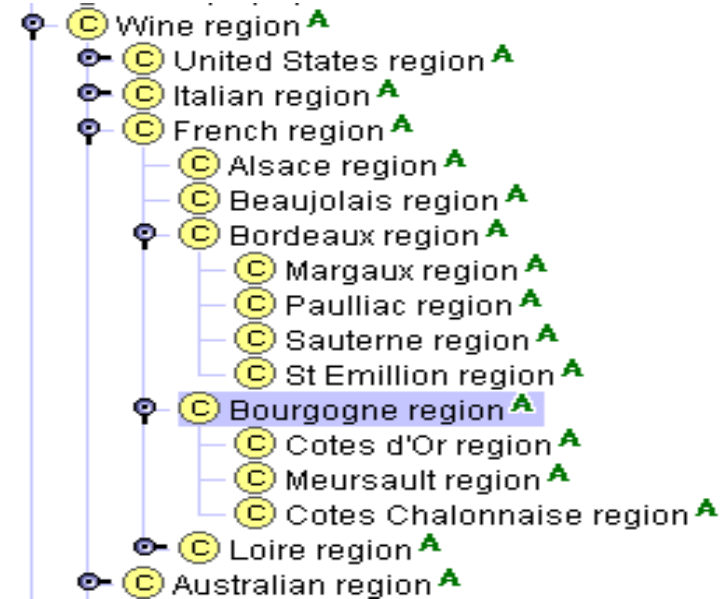
Πότε χρειάζεται να οριστεί μια νέα κλάση:

- ▶ -Συνήθως η υποκλάση μιας κλάσης α) φέρει επιπρόσθετες ιδιότητες τις οποίες η υπερκλάση δεν έχει, ή β) διαφορετικούς περιορισμούς απο την υπερκλάση, ή γ) συμμετέχει σε διαφορετικές σχέσεις απ' ότι οι υπερκλάσεις.
- Δημιουργία νέας κλάσης ή συμπλήρωση μια τιμής ενός slot;
- ▶ -Αν η προσθήκη μια τιμής σε ένα slot είναι περιοριστική για ένα slot άλλης κλάσης, τότε χρειάζεται η δημιουργία μιας κλάσης. Π.χ αν δεν δημιουργούσαμε μια κλάση White wine και συμπληρώναμε απλώς την τιμή white στο slot colour, δεν θα μπορούσαμε να συσχετίσουμε το λευκό κρασί με διάφορα φαγητά.
- ▶ -Επίσης αν η διαφοροποίηση είναι μεγάλη και σκεφτόμαστε τα αντικείμενα με διαφορετικές τιμές ως τελείως διαφορετικά αντικείμενα τότε χρειάζεται να δημιουργηθεί μια κλάση για αυτή τη διαφοροποίηση. Π.χ.η δημιουργία των κλάσεων White Merlot και Red Merlot είναι διαφορετική απ' τη δημιουργία μιας κλάσης Merlot wines, αφού έχουν μεγάλες διαφορές (διαφορετικά σταφύλλια κτλ)
- ▶ -Μια κλάση στην οποία ανήκει ένα ανεξάρτητο στιγμιότυπο δεν πρέπει να αλλάζει συχνά. π.χ. το Chilled wine δεν χρειάζεται να είναι κλάση της οντολογίας, αλλά χαρακτηριστικό της κλάσης Wine. Με τον τρόπο αυτό μπορεί να διαγραφεί σαν χαρακτηριστικό χωρίς να αλλάξει κάτι.

# Ορισμός κλάσεων και ιεραρχία κλάσεων

Στιγμιότυπο ή κλάση;

- -Τα ανεξάρτητα στιγμιότυπα αποτελούν τις πιο συγκεκριμένες δομές μιας οντολογίας(το τελευταίο επίπεδο).
- -Αν η θεματολογία ακολουθεί φυσική ιεραρχία, τότε τα στιγμιότυπα μπορούν να αναπαρασταθούν ως κλάσεις. Π.χ. το Bourgogne region είναι στιγμιότυπο της κλάσης French region. Αν θέλουμε όμως να πούμε ότι το Cotes d'Or region ανήκει στο Bourgogne region τότε το Bourgogne region θα πρέπει να να είναι κλάση. Αν γίνει αυτό όμως ίσως φανεί αυθαίρετο οπότε το protege δίνει τη δυνατότητα να οριστούν κλάσεις ως abstract.



# Ορισμός κλάσεων και ιεραρχία κλάσεων

Πότε ο καθορισμός μιας οντολογίας είναι επαρκής;

- ▶ -Η οντολογία δεν πρέπει να περιέχει όλες τις πιθανές πληροφορίες σχετικά με το αντικείμενο που μελετάται. Μόνο όσο χρειάζεται για να μην παρεκκλίνει κάποιος απ' τον σκοπό δημιουργίας της οντολογίας. Δεν χρειάζεται να γίνει υπερφόρτωση πληροφοριών οι οποίες σχετίζονται με το αντικείμενο της οντολογίας.
- Συμπέρασμα:
- ▶ -Αυτό που μπορεί να συμπεράνει κανείς είναι η διαφορετικότητα στον τρόπο υλοποίησης μιας οντολογίας. Δηλαδή δεν είναι ρητή η ορθότητα μιας οντολογίας, δεν ακολουθεί τόσο συγκεκριμένα βήματα, αλλά βασίζεται περισσότερο σε υποκειμενικά κριτήρια του δημιουργού, από τη στιγμή που ο ίδιος θέτει τους κανόνες για την υλοποίηση της.