

# ΑΝΑΠΑΡΑΣΤΑΣΗ ΓΝΩΣΗΣ ΣΤΟΝ ΠΑΓΚΟΣΜΙΟ ΙΣΤΟ

ΟΝΤΟΛΟΓΙΕΣ-Μέρος Ι

Ι. Χατζηλυγερούδης

**ΣΧΕΔΙΑΣΗ ΟΝΤΟΛΟΓΙΩΝ**  
**(βασισμένη 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) )**

# Οντολογία-Ορισμός

- Ρητός προσδιορισμός μιας εννοιολογίας (Explicit specification of a conceptualization).
- Η **εννοιολογία** αφορά το εννοιολογικό (αφηρημένο) μοντέλο ενός γνωστικού πεδίου (π.χ. ασθένεια), δηλ. τις έννοιες που το απαρτίζουν.
- Ο **ρητός προσδιορισμός** αφορά τις σχέσεις μεταξύ των εννοιών, τον ορισμό των ιδιοτήτων των εννοιών, τους περιορισμούς στις ιδιότητες.

# Οντολογία-Ορισμός

- Περιλαμβάνει ορισμούς βασικών εννοιών του πεδίου γνώσης και των σχέσεων μεταξύ τους.
- Ορίζει ένα κοινό λεξιλόγιο για μια κοινότητα ανθρώπων με κοινά γνωστικά ενδιαφέροντα.
- Οι οντολογίες ως εννοιολογικές μορφοποιήσεις (conceptualizations) αποτελούν προϊόντα υποκειμενικής κρίσης, οπότε το ίδιο πεδίο ενδιαφέροντος είναι δυνατόν να περιγραφεί με διαφορετικούς τρόπους.
- Οι περιγραφές γίνονται με αυστηρό (formal) τρόπο, ώστε να μπορούν να αναπαρασταθούν στον Η/Υ.

# Οντολογίες-Λόγοι δημιουργίας

## ■ Κοινή κατανόηση της δομής της πληροφορίας μεταξύ ανθρώπων ή μηχανών

(Π.χ. διαφορετικές ιστοσελίδες που βασίζονται σε μια κοινή οντολογία, δίνουν τη δυνατότητα αναζήτησης πληροφορίας από τη συνολική γνώση μέσω πρακτόρων).

## ■ Επαναχρησιμοποίηση της γνώσης ενός πεδίου

(Μια καλά ορισμένη οντολογία ενός πεδίου, μπορεί να επαναχρησιμοποιηθεί στον ορισμό της οντολογίας ενός ευρύτερου πεδίου).

## ■ Σαφής ορισμός των παραδοχών στην υλοποίηση οντολογίας ενός γνωστικού πεδίου

(Κάνει τις αλλαγές στην οντολογία ευκολότερες όταν οι παραδοχές πρέπει να αλλάξουν).

# Οντολογίες-Λόγοι δημιουργίας

- Διαχωρισμός της στατικής γνώσης ενός πεδίου από τη διαδικαστική γνώση

(Μία διαδικασία μπορεί να χρησιμοποιηθεί αυτούσια για διαφορετικές οντολογίες).

- Ανάλυση της γνώσης

(Η ανάλυση της γνώσης καθίσταται δυνατή με τον αυστηρό προσδιορισμό των εννοιών και των σχέσεών τους).

# Δομικά στοιχεία οντολογίας

- **Κλάσεις (Classes).** Οι κλάσεις αναπαριστάνουν έννοιες (concepts) του πεδίου. Μια έννοια μπορεί να είναι οτιδήποτε για το οποίο μπορεί να ειπωθεί κάτι και μπορεί να αφορά μια εργασία, μια λειτουργία, μια ενέργεια, μια ιδέα, κλπ.
- **Θυρίδες (Slots) ή Ιδιότητες (Properties) ή Ρόλοι (Roles).** Αναπαριστάνουν διάφορα χαρακτηριστικά (features/attributes) των κλάσεων.
- **Όψεις (Facets) ή Περιορισμοί (Restrictions).** Περιορισμοί που αφορούν τις ιδιότητες/ρόλους.
- **Στιγμιότυπα (Instances).** Τα στιγμιότυπα αναπαριστούν συγκεκριμένες οντότητες που ανήκουν σε κλάσεις.
- Οι κλάσεις και τα στιγμιότυπα αποτελούν τη βάση γνώσης (knowledge base) μιας εφαρμογής.

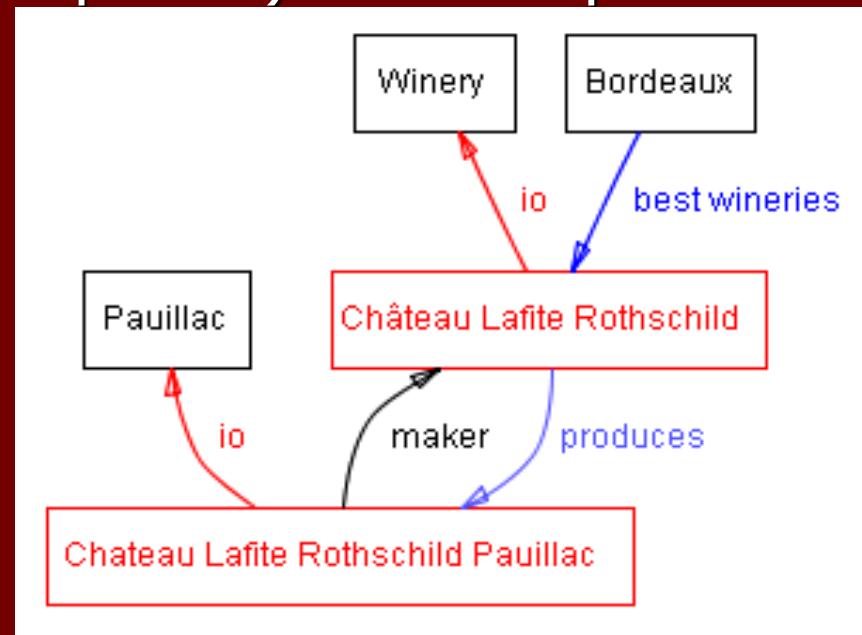
# Κλάσεις

- Τα βασικότερα στοιχεία μιας οντολογίας και αντιστοιχούν σε έννοιες, που αντιπροσωπεύουν σύνολα οντοτήτων.  
(Π.χ. η κλάση **Wine** αντιπροσωπεύει όλα τα κρασιά)
- Αποτελούνται από **ιδιότητες**, οι οποίες περιγράφονται με περιορισμούς.  
(Π.χ. η **maker** είναι μια ιδιότητα της **wine** και το **Winery** είναι περιορισμός της τιμής της **maker**-την υποχρεώνει να παίρνει ως τιμές οντότητες/στιγμιότυπα που ανήκουν στην κλάση **Winery**)
- Είναι οργανωμένες σε ιεραρχίες και σχετίζονται μεταξύ τους με τη σχέση **υποκλάση** (subclass).
- Μια υποκλάση είναι εξειδίκευση μιας κλάσης, δηλ. αντιπροσωπεύει ένα υποσύνολο οντοτήτων της κλάσης.  
(Π.χ. **RedWine** μπορεί να είναι υποκλάση της **Wine**, αντιπροσωπεύοντας τα κόκκινα κρασιά, ένα υποσύνολο των κρασιών).

# Δημιουργώντας μια οντολογία...

## ... βασικά χρειάζονται

- Ορισμός των κλάσεων.
- Ταξινόμηση των κλάσεων αυτών (υπο-κλάσεις)
- Καθορισμός των ιδιοτήτων (properties) 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".

# Βασικοί Κανόνες Σχεδίασης Οντολογιών

- **Κανόνας 1.** Δεν υπάρχει ένας «σωστός» τρόπος μοντελοποίήσης ενός πεδίου γνώσης-πάντα υπάρχουν εναλλακτικοί «σωστοί» τρόποι. Η καλύτερη λύση εξαρτάται σχεδόν πάντα από την εφαρμογή και τις πιθανές επεκτάσεις.
- **Κανόνας 2.** Η ανάπτυξη μιας οντολογίας είναι αναγκαστικά μια επαναληπτική διαδικασία: απαιτείται συχνή επανασχεδίαση.
- **Κανόνας 3.** Οι έννοιες στην οντολογία πρέπει να σχετίζονται με οντότητες (φυσικές ή λογικές) και σχέσεις μεταξύ τους στο γνωστικό πεδίο της εφαρμογής. Αυτά είναι πιο πιθανό να είναι ουσιαστικά (οντότητες) ή ρήματα (σχέσεις) σε προτάσεις που περιγράφουν το γνωστικό πεδίο.

# Βήματα για τη δημιουργία μιας οντολογίας (1)

- 1: Καθορισμός πεδίου και σκοπός της οντολογίας (εμβέλεια)
  - 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 συντηρήσει την οντολογία αυτή;

## Οντολογία κρασιών

- Προτάσεις καλών συνδυασμών κρασιού-φαγητού
- Έννοιες: κρασιά, φαγητά, καλοί συνδυασμοί (όχι έννοιες διαχείρισης αποθήκης ή κόστους)

# Βήματα για τη δημιουργία μιας οντολογίας (2)

- 2: Δυνατότητα επαναχρησιμοποίησης υπάρχουσας οντολογίας (επαναχρησιμοποίηση)
  - Έρευνα για ύπαρξη οντολογίας που σχετίζεται με την υπό δημιουργία.
  - Χρήση βιβλιοθηκών επαναχρησιμοποιούμενων οντολογιών.
  - Ενσωμάτωση υπάρχουσας οντολογίας στην ύπο δημιουργία

## Οντολογία κρασιών

Για να δημιουργήσουμε την οντολογία wines, αν υπάρχει έτοιμη μια οντολογία French wines, μπορούμε χρησιμοποιήσουμε τη βάση γνώση της ως ξεκίνημα. 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 από το άλλο.

# Βήματα για τη δημιουργία μιας οντολογίας (3)

## ■ 3: Απαρίθμηση των στοιχείων της οντολογίας.

-Καταγραφή όλων των όρων που μπορεί να είναι χρήσιμοι για την υπό κατασκευή δημιουργία, χωρίς σκέψεις καταλληλότητας. Στα επόμενα βήματα θα επιλεγούν και θα ταξινομηθούν κατάλληλα.

## Οντολογία κρασιών

Για την οντολογία αυτή μπορούμε να καταγράψουμε όρους όπως, wine, winery, grape, location, colour, sugar rate, food accompaniment.

# Βήματα για τη δημιουργία μιας οντολογίας (4)

- **Βήμα 4: Καθορισμός των κλάσεων και της ιεραρχίας κλάσεων**  
-Προσδιορίζουμε ποιες έννοιες αποτελούν κλάσεις και ποιες ιδιότητες. Επίσης, προσδιορίζουμε τις σχέσεις «υποκλάση» μεταξύ των κλάσεων και στη συνέχεια αναπτύσσουμε την ιεραρχία των κλάσεων με βάση μια από τις παρακάτω μεθόδους:
  - top-down.** Ξεκινούμε από την γενικότερη κλάση (που χαρακτηρίζει την οντολογία) και κάποιες βασικές υποκλάσεις της. Στη συνέχεια εξειδικεύουμε τις υποκλάσεις κ.ο.κ.
  - bottom-up.** Ορίζουμε τις κλάσεις του τελευταίου επιπέδου (τα φύλλα του δέντρου της ιεραρχίας), τις οποίες οργανώνουμε ως υποκλάσεις κάποιων κλάσεων και αυτές άλλων πιο πάνω κ.ο.κ.
  - mixed.** Συνδυασμός των δύο παραπάνω. Κάποιος μπορεί να ξεκινήσει με top-down και να συνεχίσει με bottom-up ή αντίστροφα.

# Βήματα για τη δημιουργία μιας οντολογίας (5)

## Οντολογία κρασιών

-Π.χ. ορίζουμε τα Wine, Winery ως κλάσεις και τα colour, grape, location ως ιδιότητες.

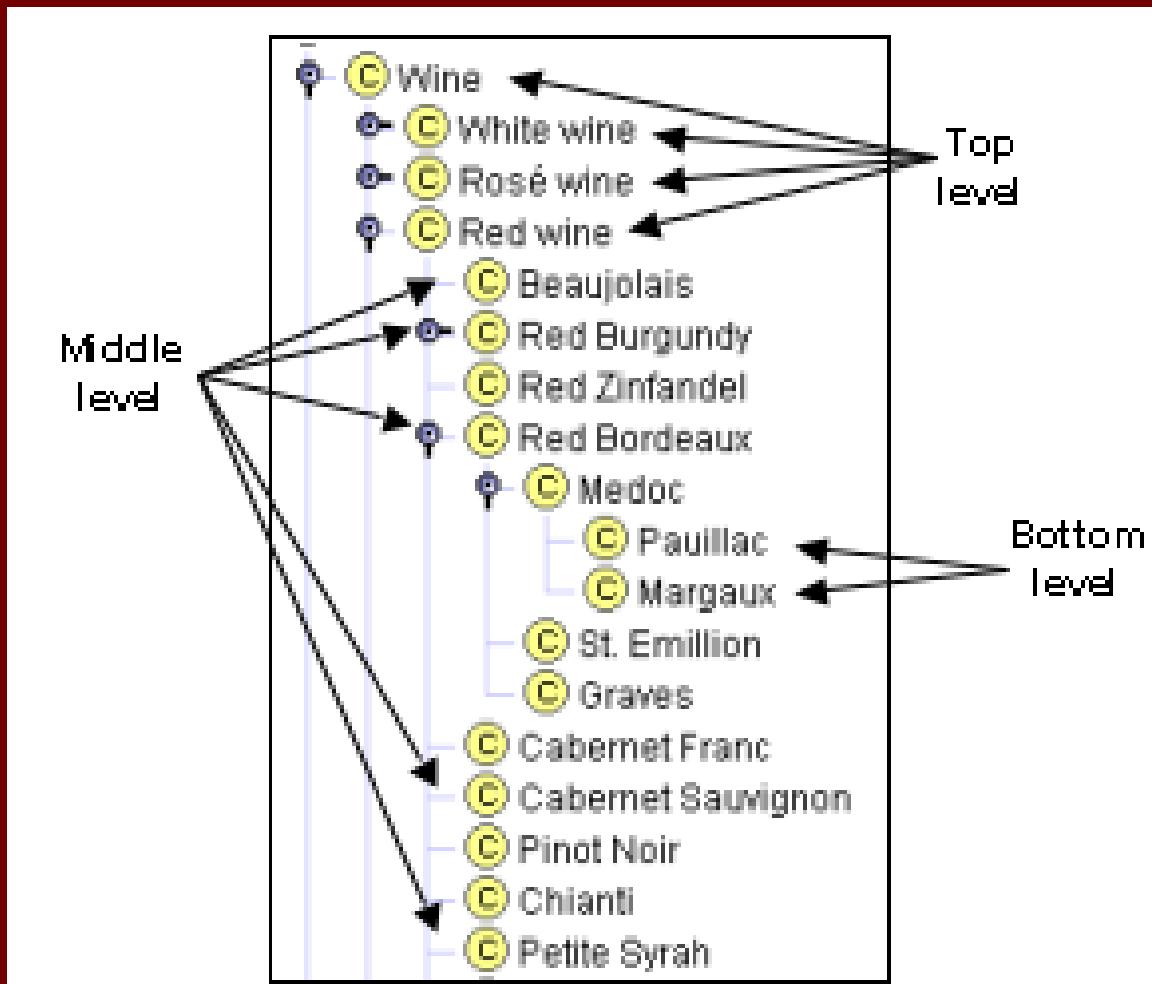
-**top-down**. Ξεκινούμε από την Wine και θεωρούμε ως υποκλάσεις τις Red wine, White wine και Rose wine. Στη συνέχεια, προχωρούμε με τις Red Burgundy, Cabernet Sauvignon ως υποκλάσεις της Red wine κ.ο.κ.

-**bottom-up**. Ξεκινάμε π.χ. από τις κλάσεις Pauillac και Margeux, ως τερματικές κλάσεις, και τις ορίζουμε ως υποκλάσεις της Medoc, την οποία, μαζί με τις St. Emillion και Graves, ορίζουμε ως υποκλάσεις της Red Bordeaux.

-**mixed**. Ξεκινάμε από κάποιες ενδιάμεσες κλάσεις, π.χ. Red Burgundy, Red Bordaeux, και προχωράμε bottom-up ορίζοντας τες ως υποκλάσεις της Red wine, και συνεχίζουμε top-down δημιουργώντας τις St. Emillion και Medoc ως υποκλάσεις της Red Bordaeux κ.ο.κ.

# Βήματα για τη δημιουργία μιας οντολογίας (6)

- Μέρος οντολογίας wine στο protege (εργαλείο ανάπτυξης οντολογιών)



# Βήματα για τη δημιουργία μιας οντολογίας (7)

- 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 (8)

## ■ 5: Καθορισμός των ιδιοτήτων των κλάσεων (slots)

-Γίνεται προσδιορισμός των ιδιοτήτων (properties) και κατανέμονται στις κλάσεις της ιεραρχίας.

-Οι κλάσεις από μόνες τους δεν μπορούν να δώσουν αρκετές πληροφορίες. Αυτό γίνεται με τον ορισμό των ιδιοτήτων τους.

-Σε μια οντολογία υπάρχουν διάφορα είδη ιδιοτήτων:

\* «ενδογενείς» ή «φυσικές» ιδιότητες (π.χ. flavor)

\* «εξωγενείς» ιδιότητες (π.χ. name, area),

\* «μέρη», αν οι οντότητες μιας κλάσης είναι δομημένες (π.χ. τα στάδια ενός γεύματος)

\* «σχέσεις» μεταξύ ανεξάρτητων μελών της κλάσης και άλλων οντοτήτων (π.χ maker της Wine).

# Βήματα για τη δημιουργία μιας οντολογίας (9)

- 6: Καθορισμός όψεων (facets) των ιδιοτήτων (slots) των κλάσεων
  - Γίνεται περιγραφή των όψεων, δηλ. των περιορισμών στις τιμές των ιδιοτήτων των κλάσεων. Οι όψεις αφορούν περιορισμούς όπως, αριθμός τιμών (cardinality), τύπος τιμών (type) κλπ. Κάθε όψη έχει ένα γέμισμα ή τιμή (filler).
  - Συνήθεις όψεις:
    - Slot cardinality.** Ορίζει πόσες τιμές μπορεί να έχει μια ιδιότητα. Π.χ η ιδιότητα body της κλάσης Wine μπορεί να πάρει μόνο μια τιμή (single cardinality slot) αφού ένα κρασί έχει ένα τύπο ως προς το πόσο δυνατό είναι.
    - Slot-value type.** Ορίζει τον τύπο των τιμών μιας ιδιότητας, π.χ String, Number, Boolean, Enumerated (εδώ απαριθμούμε τις τιμές: π.χ. για την flavor μπορεί να είναι strong, moderate, delicate)
    - Instance-type slot:** Π.χ. η ιδιότητα produces, που είναι τύπου instance, της κλάσης Winery μπορεί να έχει ως τιμές στιγμιότυπα της κλάσης Wine.

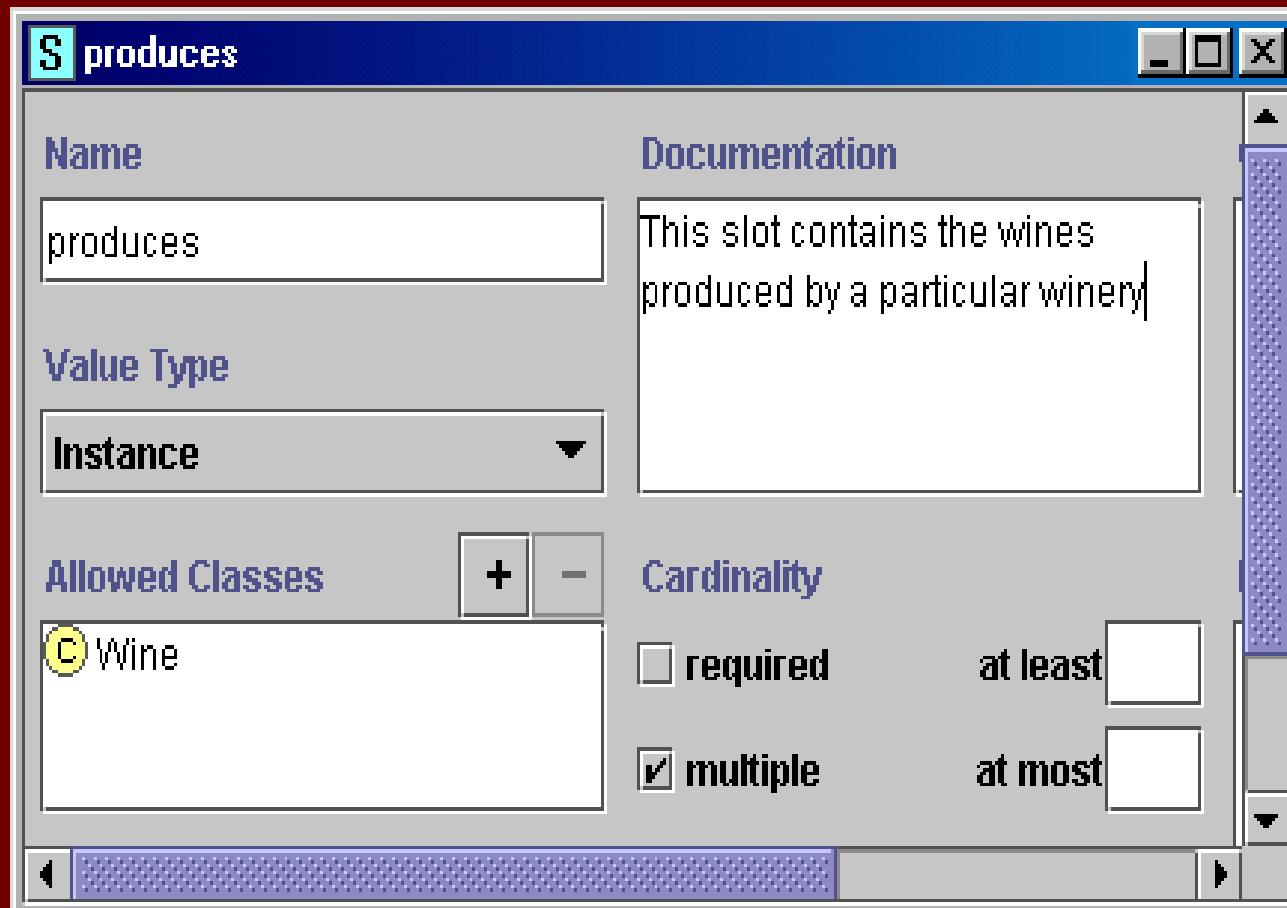
# Βήματα για τη δημιουργία μιας οντολογίας (10)

- Οι ιδιότητες (slots) της κλάσης Wine και οι όψεις τους (facets), 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 <span style="color: #00AEEF;">I</span>	Instance	single	classes={Winery}						
S	name	String	single							
S	sugar	Symbol	single	allowed-values={DRY,SWEET,OFF-DRY}						

# Βήματα για τη δημιουργία μιας οντολογίας (11)

- Το παρακάτω σχήμα δείχνει την σχέση της ιδιότητας produces με την κλάση Winery.



# Βήματα για τη δημιουργία μιας οντολογίας (12)

**Πεδίο ορισμού (domain) και Πεδίο τιμών (range) ιδιότητας**

Οι κλάσεις ως τιμές (fillers) σε ιδιότητες τύπου instance αποτελούν το **πεδίο τιμών (range)** της ιδιότητας.

Οι κλάσεις στις οποίες προσαρτάται (ή ανήκει) μια ιδιότητα αποτελούν το **πεδίο ορισμού (domain)** της ιδιότητας.

Γενικός κανόνας: Για να ορίσουμε τα πεδία ορισμού και τιμών, βρίσκουμε τις πιο γενικές κλάσεις οι οποίες μπορούν να τα αποτελούν.

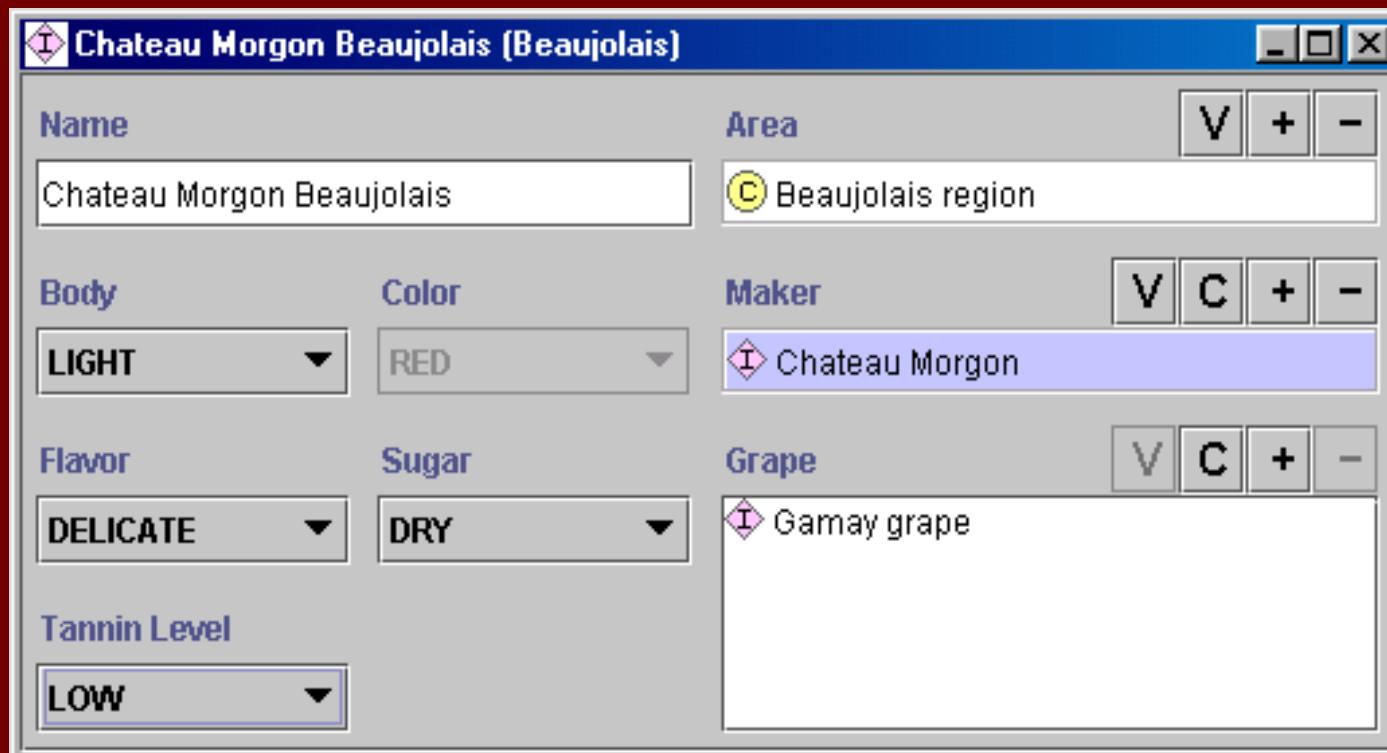
## Οντολογία κρασιών

Η produces είναι μια ιδιότητα τύπου στιγμιοτύπου της κλάσης Winery με τιμή Wine. Η Wine (δηλ. τα στιγμιότυπα της Wine) αποτελεί το πεδίο τιμών της produces, ενώ η Winery αποτελεί το πεδίο ορισμού της produces.

# Βήματα για τη δημιουργία μιας οντολογίας (13)

## ■ 7: Δημιουργία στιγμιοτύπων

-Για την δημιουργία ενός στιγμιοτύπου μιας κλάσης: (1)επιλέγουμε την κλάση, (2) δημιουργούμε ένα (κενό) στιγμιότυπο της κλάσης, (3) συμπληρώνουμε τις τιμές των ιδιοτήτων του.



# Ορισμός και ιεραρχία κλάσεων (1)

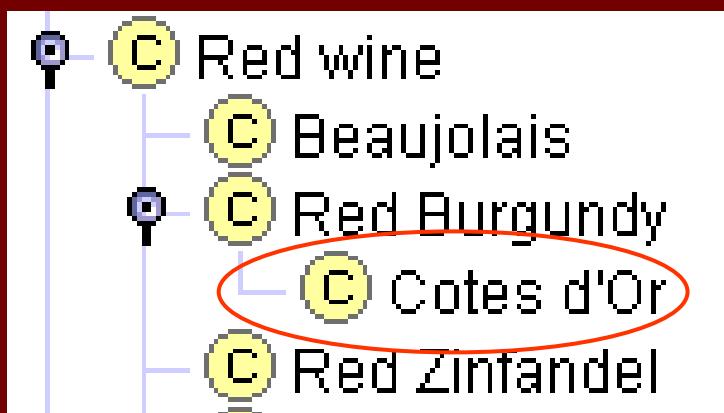
## ■ Σημεία ελέγχου συνέπειας ιεραρχίας κλάσεων

- Μια οντολογία βασίζεται στη σχέση “is-a”: Μια κλάση A είναι υποκλάση της B, αν οποιοδήποτε στιγμιότυπο της A είναι και στιγμιότυπο της B. Δηλαδή είναι μια σχέση συνόλου-υποσυνόλου.
- Η σχέση “part-of” δεν αποτυπώνεται σαν σχέση ιεραρχίας της οντολογίας, αλλά ως ιδιότητα.
- Η σχέση «υποκλάση» είναι μεταβατική: Αν η κλάση B είναι υποκλάση της A και η Γ είναι υποκλάση της B, τότε και η Γ είναι υποκλάση της B.
- Συνιστάται η χρήση ονομάτων κλάσεων στον ενικό.
- Αποφυγή χρήσης διαφορετικών ονομάτων για την ίδια κλάση (π.χ. Shrimp-Prawn).
- Αποφυγή κύκλων στις κλάσεις: Αυτό σημαίνει η κλάση A να έχει μια υποκλάση B και ταυτόχρονα η B να έχει υποκλάση την A.

# Ορισμός και ιεραρχία κλάσεων (2)

## ■ Ανάλυση αδελφικών (*sibling*) κλάσεων

- Αδελφικές (*Siblings*) είναι κλάσεις, οι οποίες είναι υποκλάσεις των ίδιων κλάσεων και πρέπει να βρίσκονται στο ίδιο επίπεδο γενικότητας.
- Εάν μια κλάση έχει ένα μόνο μια υποκλάση (δηλ. δεν υπάρχουν αδελφικές κλάσεις), τότε υπάρχει πρόβλημα μοντελοποίησης ή η ιεραρχία είναι δεν είναι πλήρης.
- Αν μια κλάση έχει πάνω από 12 αδελφικές υποκλάσεις, τότε χρειάζεται περαιτέρω (ενδιάμεση) κατηγοριοποίηση.



# Ορισμός και ιεραρχία κλάσεων (3)

## **Πότε χρειάζεται να οριστεί μια νέα (υπο)κλάση:**

Συνήθως η υποκλάση μιας κλάσης α) φέρει επιπρόσθετες ιδιότητες τις οποίες η υπερκλάση δεν έχει, ή β) διαφορετικούς περιορισμούς από την υπερκλάση, ή γ) συμμετέχει σε διαφορετικές σχέσεις απ' ότι οι υπερκλάσεις.

## **Δημιουργία νέας (υπο)κλάσης ή συμπλήρωση μια τιμής μιας ιδιότητας;**

Αν η προσθήκη μια τιμής σε μία ιδιότητα είναι περιοριστική για μία ιδιότητα άλλης κλάσης, τότε χρειάζεται η δημιουργία μιας νέας κλάσης. Π.χ αν δεν δημιουργούσαμε μια κλάση White wine και συμπληρώναμε απλώς την τιμή white στην ιδιότητα colour, δεν θα μπορούσαμε να συσχετίσουμε το λευκό κρασί με διάφορα φαγητά.

# Ορισμός και ιεραρχία κλάσεων (4)

Αν μια διάκριση μεταξύ εννοιών στο γνωστικό πεδίο είναι σημαντική και σκεφτόμαστε τα αντικείμενα με διαφορετικές τιμές για την διάκριση αυτή ως αντικείμενα διαφορετικού είδους, τότε χρειάζεται να δημιουργηθεί μια κλάση για αυτή τη διάκριση. 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 σταφύλια κτλ)

Μια κλάση στην οποία ανήκει ένα στιγμιότυπο δεν πρέπει να αλλάζει συχνά. Π.χ. το Chilled wine δεν χρειάζεται να είναι κλάση της οντολογίας, αλλά χαρακτηριστικό της κλάσης Wine. Ο χαρακτηρισμός Chilled wine για τα μπουκάλια κρασιού δεν είναι μόνιμος ή με άλλα λόγια η κατάσταση chilled για ένα μπουκάλι κρασί δεν είναι μόνιμη, οπότε ένα κρασί θα μεταπηδά από την κατάσταση chilled στην normal.

# Ορισμός και ιεραρχία κλάσεων (5)

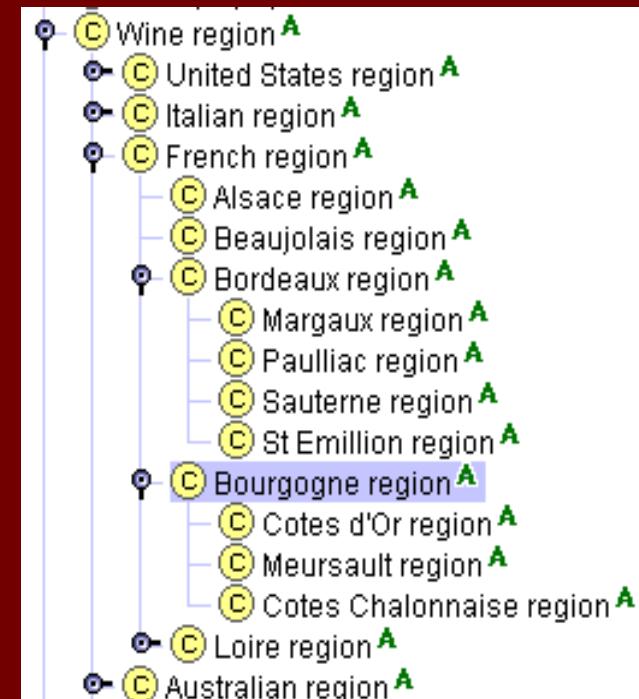
## Στιγμιότυπο ή κλάση;

Μία έννοια πολλές φορές μπορεί να μοντελοποιηθεί είτε ως κλάση είτε ως στιγμιότυπο.

Π.χ. υπάρχει ένα συγκεκριμένο κρασί το Sterling Vineyards Merlot, το οποίο μπορεί να μοντελοποιηθεί είτε ως στιγμιότυπο (π.χ. της κλάσης Merlot), είτε ως κλάση (υποκλάση της κλάσης Merlot). Αυτό εξαρτάται από τη χρήση της οντολογίας.

Αν έννοιες σχηματίζουν μια φυσική ιεραρχία, τότε αυτές πρέπει να αναπαρασταθούν ως κλάσεις.

Π.χ. η ιεραρχία των οινοπαραγωγικών περιοχών (βλ. σχήμα). Συνήθως αυτές οι κλάσεις ορίζονται ως **αφηρημένες** (abstract), δηλ. κλάσεις χωρίς στιγμιότυπα.



# Ορισμός και ιεραρχία κλάσεων (6)

Πότε ο καθορισμός μιας οντολογίας είναι επαρκής;  
Δύο κανόνες:

- Η οντολογία δεν πρέπει να περιέχει όλες τις πιθανές πληροφορίες για το γνωστικό πεδίο: δεν χρειάζεται να εξειδικεύει (ή να γενικεύει) περισσότερο από όσο χρειάζεται για την εφαρμογή (το πολύ ένα επιπλέον επίπεδο).
  
- Η οντολογία δεν πρέπει να περιέχει όλες τις πιθανές ιδιότητες και διακρίσεις μεταξύ των κλάσεων στην ιεραρχία.