

臨床医学オントロジー研究開発の 必要性・現状と課題

東京大学 大学院医学系研究科

今井 健

Project Leader 大江和彦

概要

1. はじめに
2. 疾患概念の記述の概要
3. 人体解剖構造の記述の概要
4. 臨床医学オントロジーの活用展望と課題

はじめに

- 背景
 - 電子的に蓄積される情報の増加
 - 新たな医学的知識の発見, 類似症例検索, 医療事故防止支援へ2次利用
 - これまでは用語の標準化、情報モデルの標準化 (診療情報記述, 交換, etc.)
 - 数々の医学概念 (疾患, 症状, 部位, 検査, 薬剤) を計算機で扱い意味処理を行うための知識ベースが必要
- 1つの方式として臨床医学オントロジーの構築に取り組んできた (臨床医学概念の定義を、概念間の意味関係を用いて形式的に記述した体系)
- 過去のオントロジー開発
 - SNOMED-CT, GALEN, FMA, GO
 - オントロジー工学から見たときの整合性に問題あり
- 2007年度より厚労省受託研究として東京大学・千葉大学・大阪大学のグループで、我が国独自フレームワークのオントロジー開発を開始。

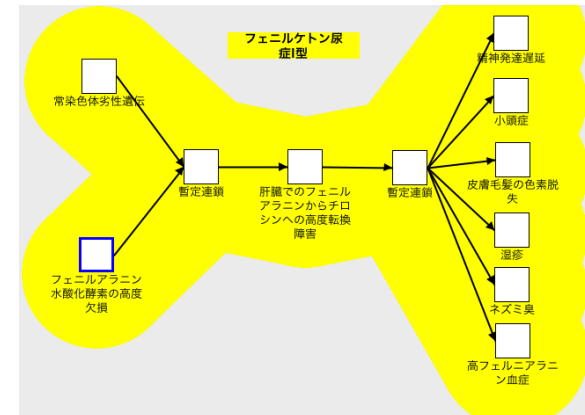
疾患概念の記述の概要

- 基本方針
 - バランス異常の導入
 - 状態連鎖としての疾患
- 汎用状態連鎖と因果推論
- 記述できていないもの
- 現状規模

基本方針

- 疾患とは何か

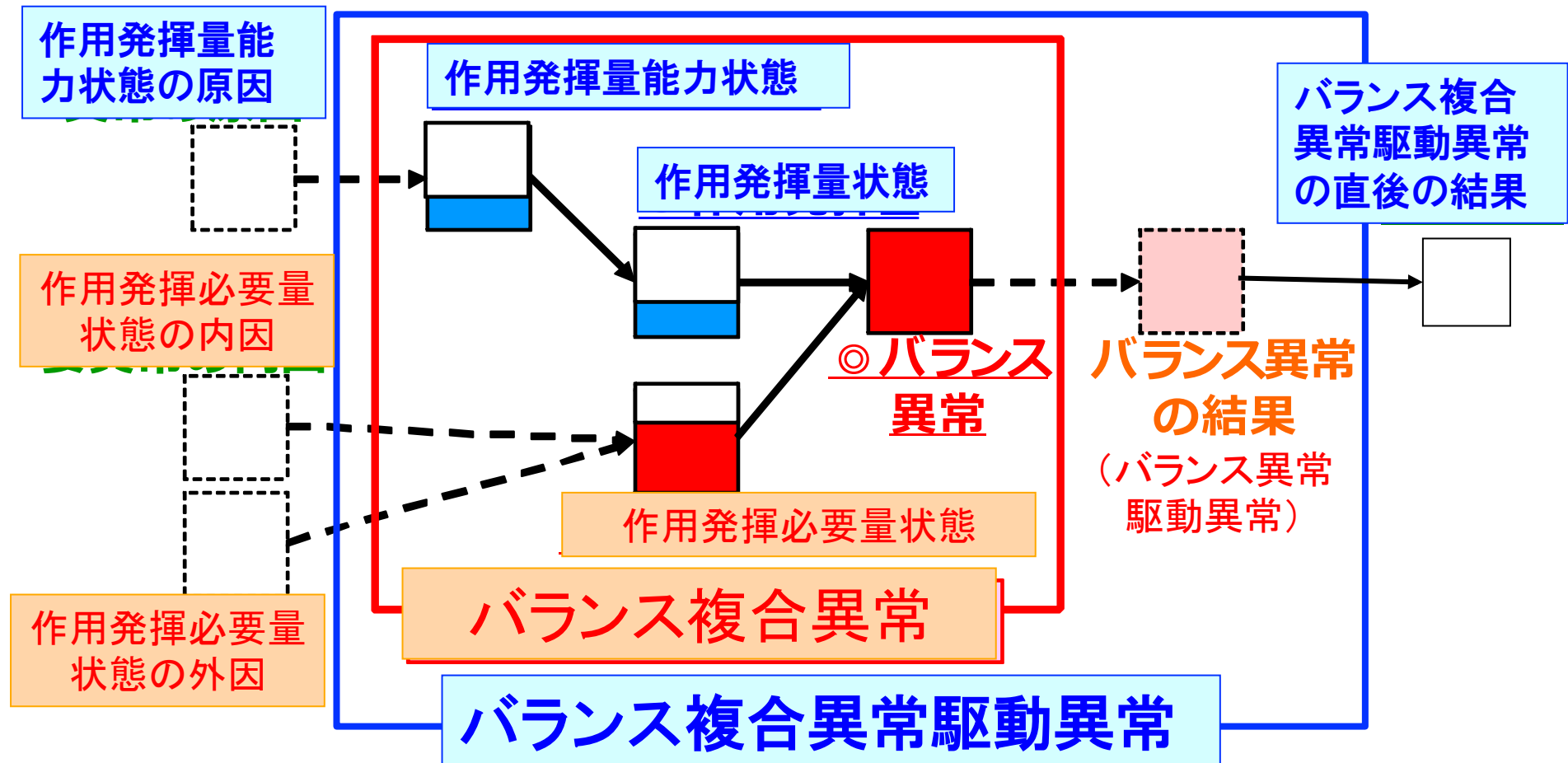
- 心身に発生した何らかの異常状態？
(ex) 糖尿病 = 持続的な高血糖状態
- 血糖値を良好にコントロールされている患者は糖尿病では無いのか？



- 本研究での記述フレームワーク

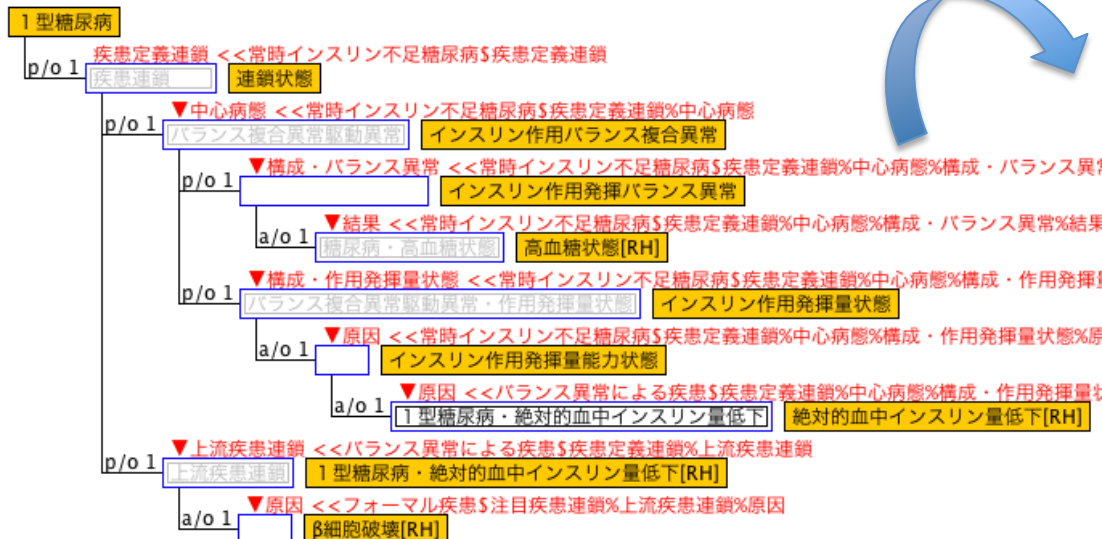
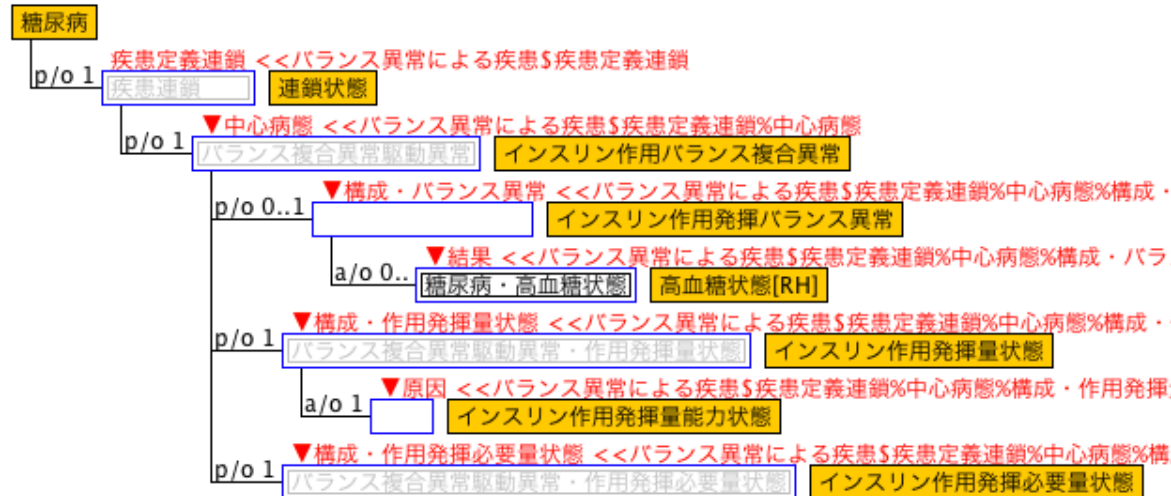
- (1) バランス異常 (作用発揮必要量と発揮量のバランスが本質的に崩れている状態) を導入した上で、
- (2) 原因と途中経過を含めた「一連の状態変化の連鎖」として記述

(1) バランス異常の導入



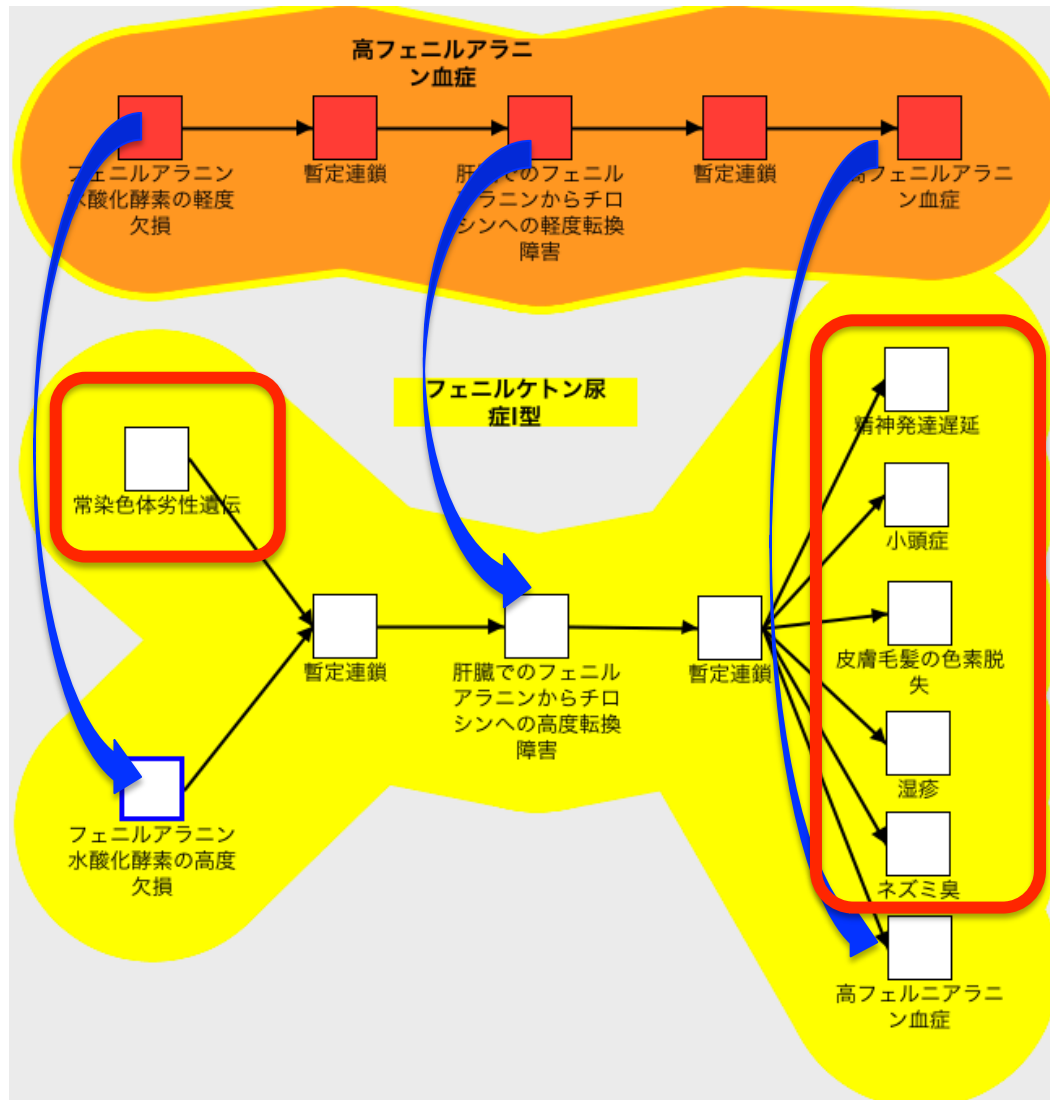
(1) バランス異常の導入

糖尿病の定義 & 1型糖尿病の定義



介入により血糖値をコントロールされている糖尿病患者でも、本質的に「インスリン需給のバランス異常を起こす状態」であるため糖尿病であるということが自然に記述できる

(2) 疾患概念の状態因果連鎖



高フェニルアラニン血症

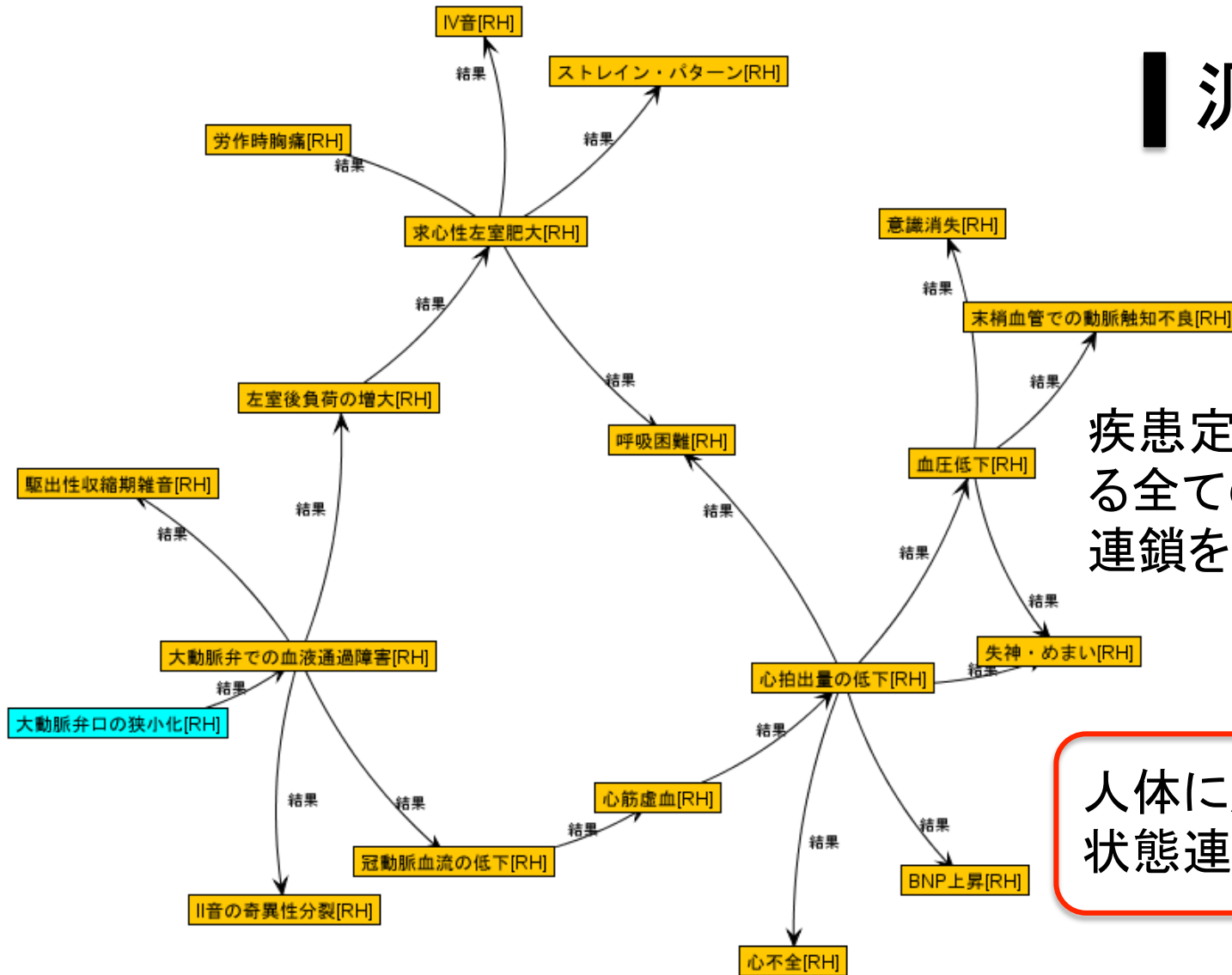
↑ IS-A (汎化特化) 関係

フェニルケトン尿症 I型

- 1) 状態連鎖内の対応するノードが、同じか、詳細化されている
- 2) 下位概念では状態連鎖に新規ノードが追加されている

汎化・特化関係が理解しやすくなる

汎用連鎖



疾患定義において現れる全ての異常状態の連鎖を集めたもの



人体に起こりえる異常状態連鎖全体のセット

汎用連鎖



複数の疾患にまたがった
因果関係推論を行う仕組み



将来的な知識発見に大いに
寄与すると考えられる


まだ記述できていないもの

- **現在の医学で、原因や病態生理が分かっていないもの**
 - どういうメカニズムで疾患が起こっているか分からないが、典型的な症状・所見の組み合わせに疾患名が付いているもの (症候群など)
 - 途中に「暫定連鎖ノード」を仮に挿入して記述
- **状態連鎖における各異常状態ノードが「どう観測されるか」**
 - (ex) 糖尿病における高血糖状態
 - 検査値での定義は数年おきに変更されている
 - オントロジー中では普遍的な概念を定義。観測時でのクライテリア (空腹時血糖値 $\geq 126\text{mg/dL}$) は別レイヤーで記述する方針
 - 2010年, 1999年における糖尿病診断基準が異なっても、オントロジーは変更を受けない。
クライテリアの種類で個々の患者の糖尿病を区別。

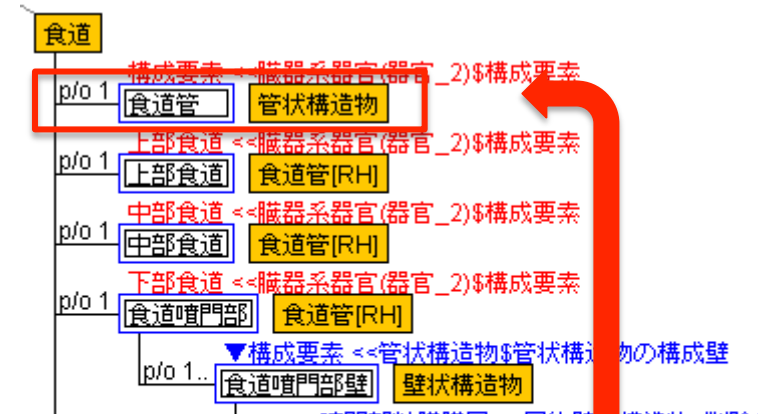
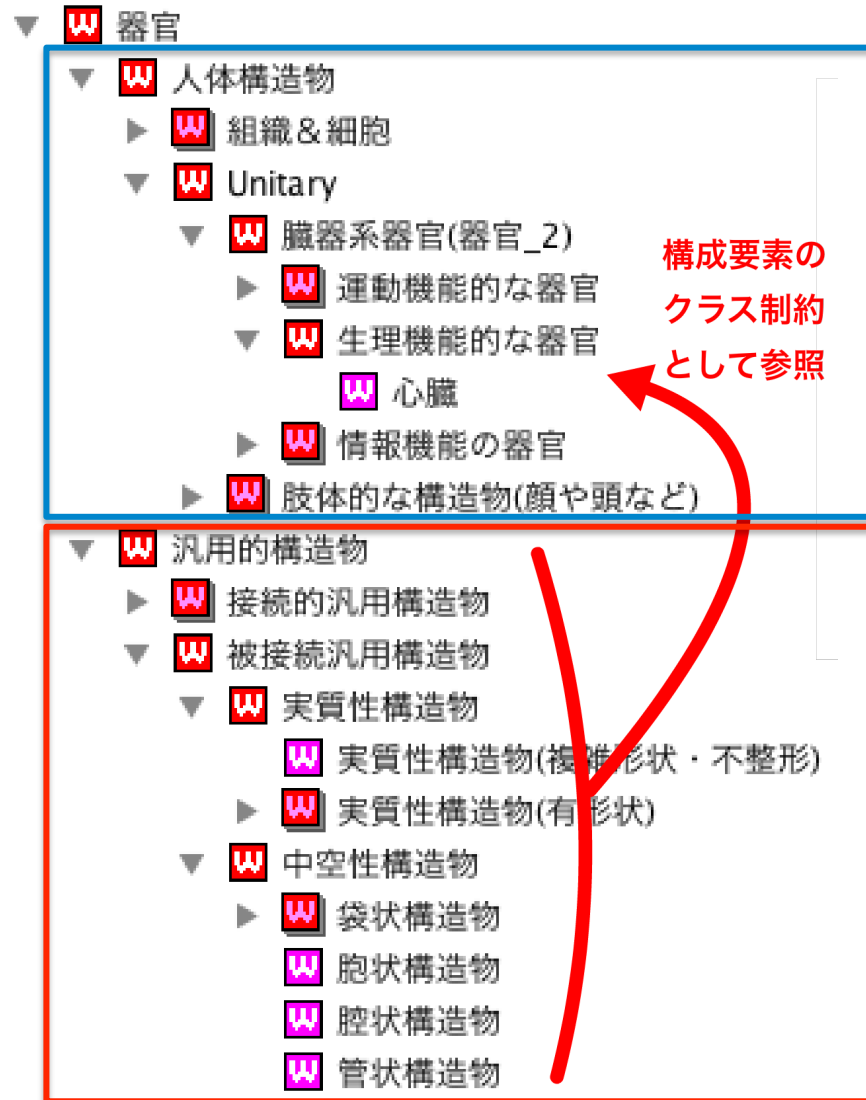
疾患オントロジーの現状規模

- 12診療科の医師によりテンプレート入力されたデータをオントロジー形式に変換
 - アレルギー・リウマチ内科, 眼科, 血液内科, 呼吸器内科, 耳鼻科, 循環器内科, 小児科, 神経内科, 腎臓内分泌内科, 整形外科, 糖尿病代謝内科, 皮膚科
- 計6,051疾患概念
 - 疾患を含む全構成概念数: 約15,000
 - 全スロット数: 約60,000
- 現在記述精度向上のため見直し作業中

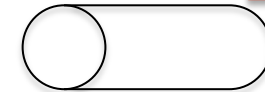
人体解剖構造の記述の概要

- 
- ・共通性と特殊性の表現
 - ・全体性と部分性の表現
 - ・接続の表現
 - ・現状規模

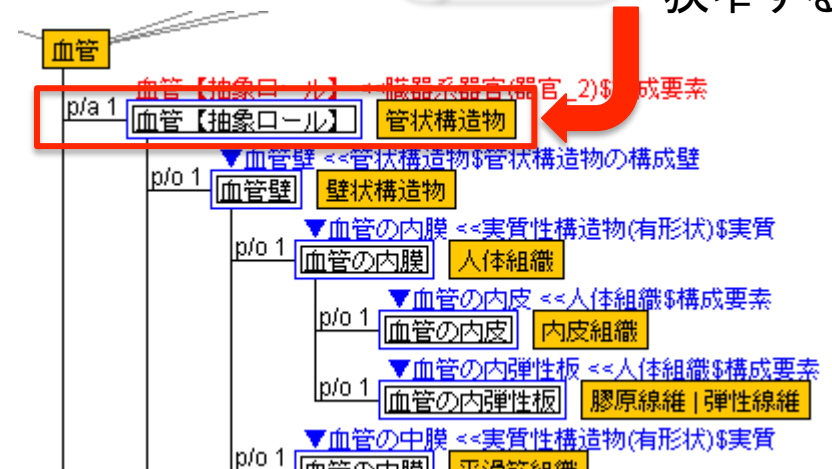
共通性と特殊性の表現



管状構造物



閉塞する
狭窄する



全体性と部分性の表現

- 「1つの解剖構造物を、より詳細な構成部品の集合として記述する」ときの基本的な意味関係は「部分-全体関係」(part-of)
- しかし考慮すべき問題が...
 - **Q1:** 部分に分解する時の複数の観点をどうするか?
「胃は袋と壁からなる」
vs 「胃は胃体部, 胃底部, 噴門部, 幽門部からなる」
 - **Q2:** 胃角は胃のある特定領域・部分に名前が付いているだけで「胃の構成部品」とは言い難いが、胃の一部であることは間違いない。構成部品としての「part-of」を用いて記述して良いか?

1種類の part-of 関係だけでは対応できない !



目的に応じて使い分ける**4種類の part-of 関係**を導入

【p/o】

【通常のp/oでとらえた人体構造の記述】

- 部品的なもの【part】
 - 胃体部, 噴門部, 左心房, 右心室など
部分・部品をさす通常のpart

【p/o-w】

【別視点でとらえた人体構造の記述】

- 別の視点から見た全体をさすもの【whole】
 - 袋状の構造物として見た胃全体 (内腔と胃壁)
vs 入り口・中央部・出口 (噴門部, 胃体部, 幽門部)

【p/o-r】

【任意の領域に部分名称がついている人体構造の記述】

- 領域名称・部分名称的なもの【region】
 - 胃角, 小弯, 大弯, 心先など

【p/o-gc】

【コンテキスト依存構造物】

- コンテキスト依存の汎用的構造物で他の部品や部分のクラス制約として参照されるロールホルダー【gc】
 - 胃壁や心房壁など, 共通性のある構造が見られるもの
 - 心臓にあるだけの心臓壁など

【p/o】

【通常のp/oでとらえた人体構造の記述】

- 部品的なもの【part】
 - 胃体部, 噴門部, 左心房, 右心室
部分・部品をさす通常のpart

【p/o-w】

【別視点でとらえた人体構造の記述】

- 別の視点から見た全体をさすもの【view】
 - 袋状の構造物として見た胃全体
vs 入り口・中央部・出口 (噴門部)

【p/o-r】

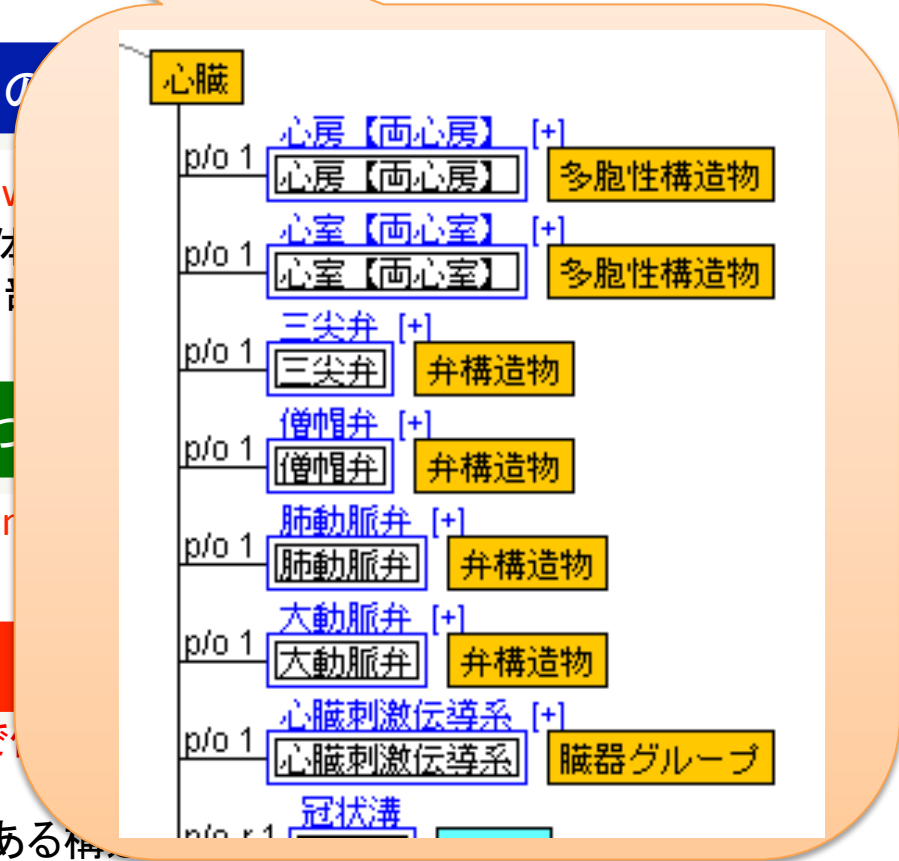
【任意の領域に部分名称がついたもの】

- 領域名称・部分名称的なもの【region】
 - 胃角, 小弯, 大弯, 心先など

【p/o-gc】

【コンテキスト依存構造物】

- コンテキスト依存の汎用的構造物で参照されるロールホルダー【gc】
 - 胃壁や心房壁など, 共通性のある構造物
 - 心臓にあるだけの心臓壁など



【p/o】 【通常のp/oでとらえた人体構造の記述】

- 部品的なもの【part】
 - 胃体部, 噴門部, 左心房, 右心室など部分・部品をさす通常のpart

【p/o-w】 【別視点でとらえた人体構造の記述】

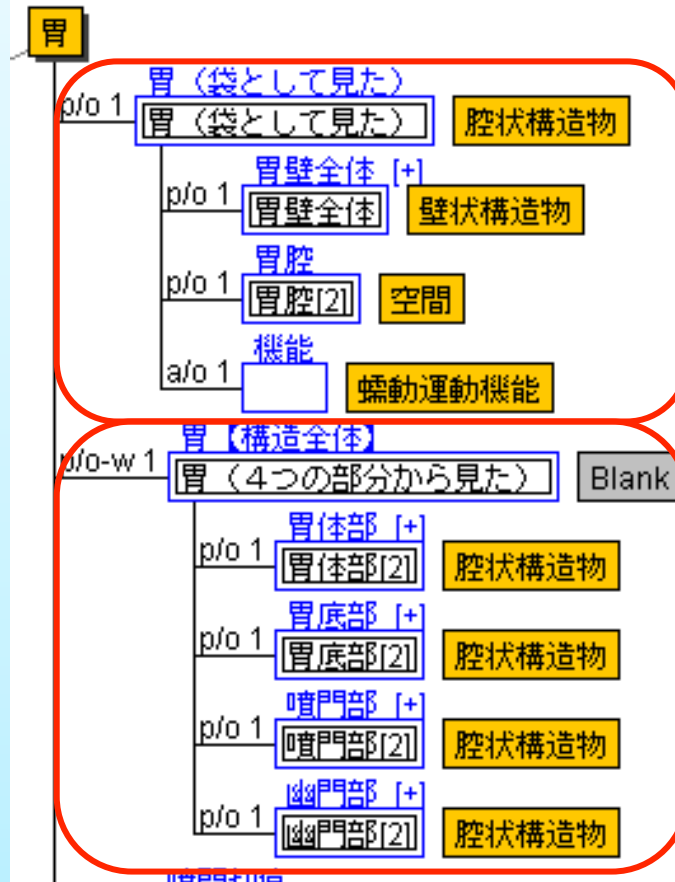
- 別の視点から見た全体をさすもの【whole】
 - 袋状の構造物として見た胃全体 (内腔) vs 入り口・中央部・出口 (噴門部, 胃体部)

【p/o-r】 【任意の領域に部分名称がついていないもの】

- 領域名称・部分名称的なもの【region】
 - 胃角, 小弯, 大弯, 心先など

【p/o-gc】 【コンテキスト依存構造物】

- コンテキスト依存の汎用的構造物で他の部分に参照されるロールホルダー【gc】
 - 胃壁や心房壁など, 共通性のある構造物
 - 心臓にあるだけの心臓壁など



複数の視点による分割を表現

【p/o】 【通常のp/oでとらえた人体構造の記述】

- 部品的なもの【part】
 - 胃体部, 噴門部, 左心房, 右心室など
部分・部品をさす通常のpart

【p/o-w】 【別視点でとらえた人体構造の記述】

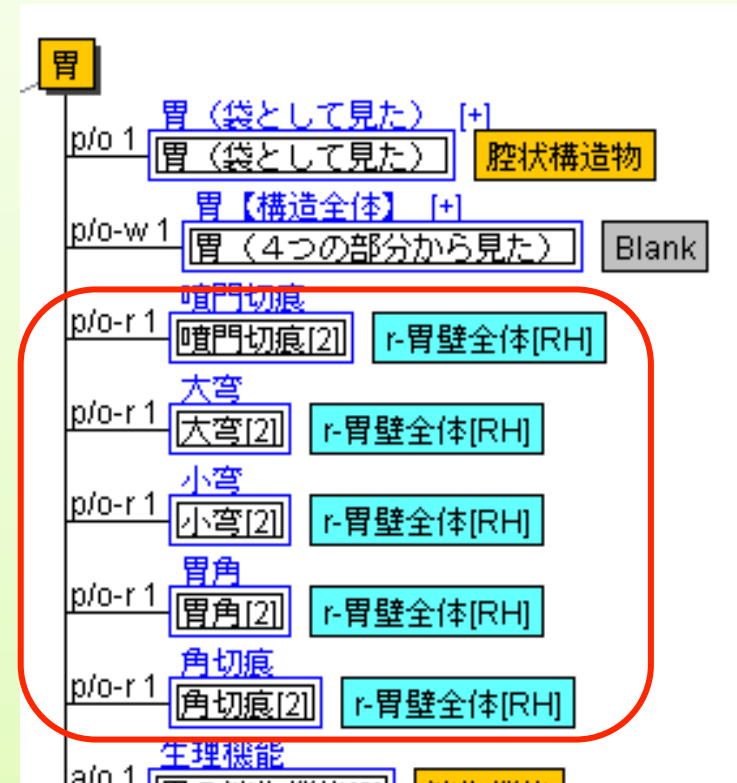
- 別の視点から見た全体をさすもの【whole】
 - 袋状の構造物として見た胃全体 (内腔)
vs 入り口・中央部・出口 (噴門部, 胃体)

【p/o-r】 【任意の領域に部分名称がついて】

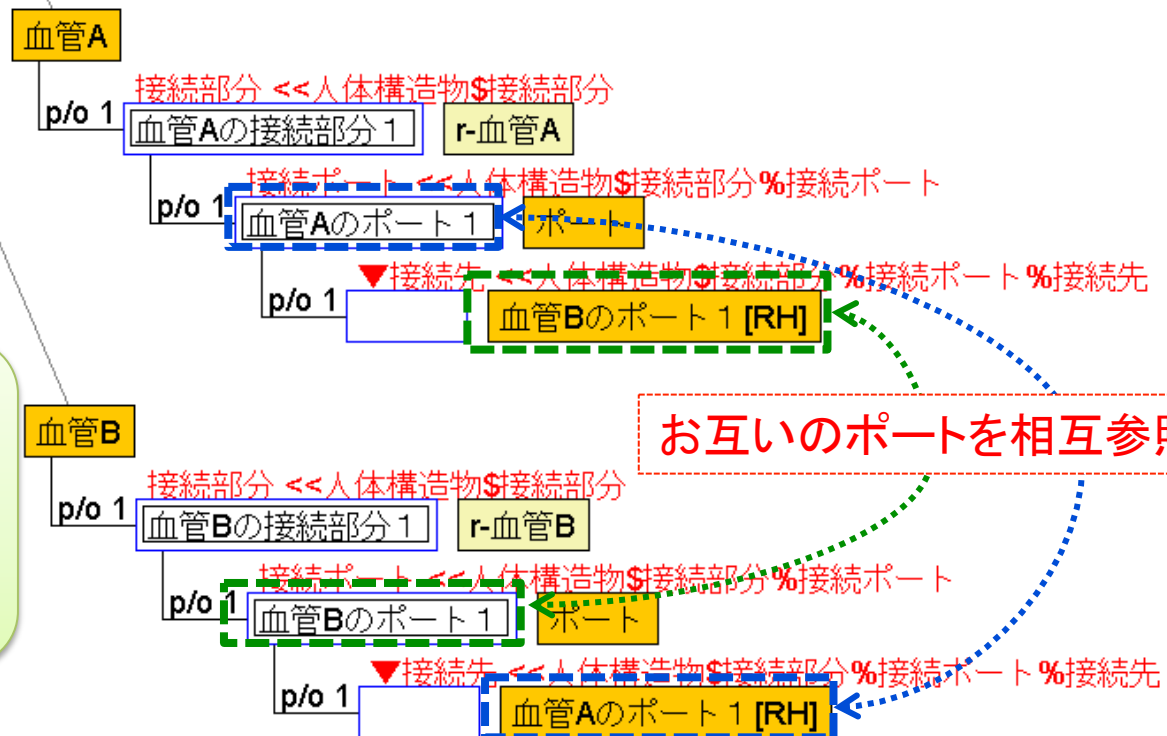
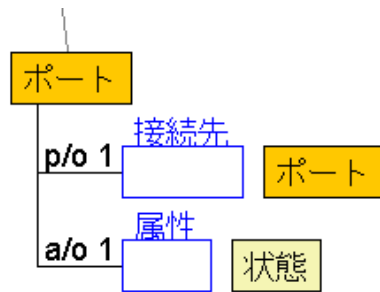
- 領域名称・部分名称的なもの【region】
 - 胃角, 小弯, 大弯, 心先など

【p/o-gc】 【コンテキスト依存構造物】

- コンテキスト依存の汎用的構造物で他の部品
参照されるロールホルダー【gc】
 - 胃壁や心房壁など, 共通性のある構造が見られるもの
 - 心臓にあるだけの心臓壁など



接続の表現

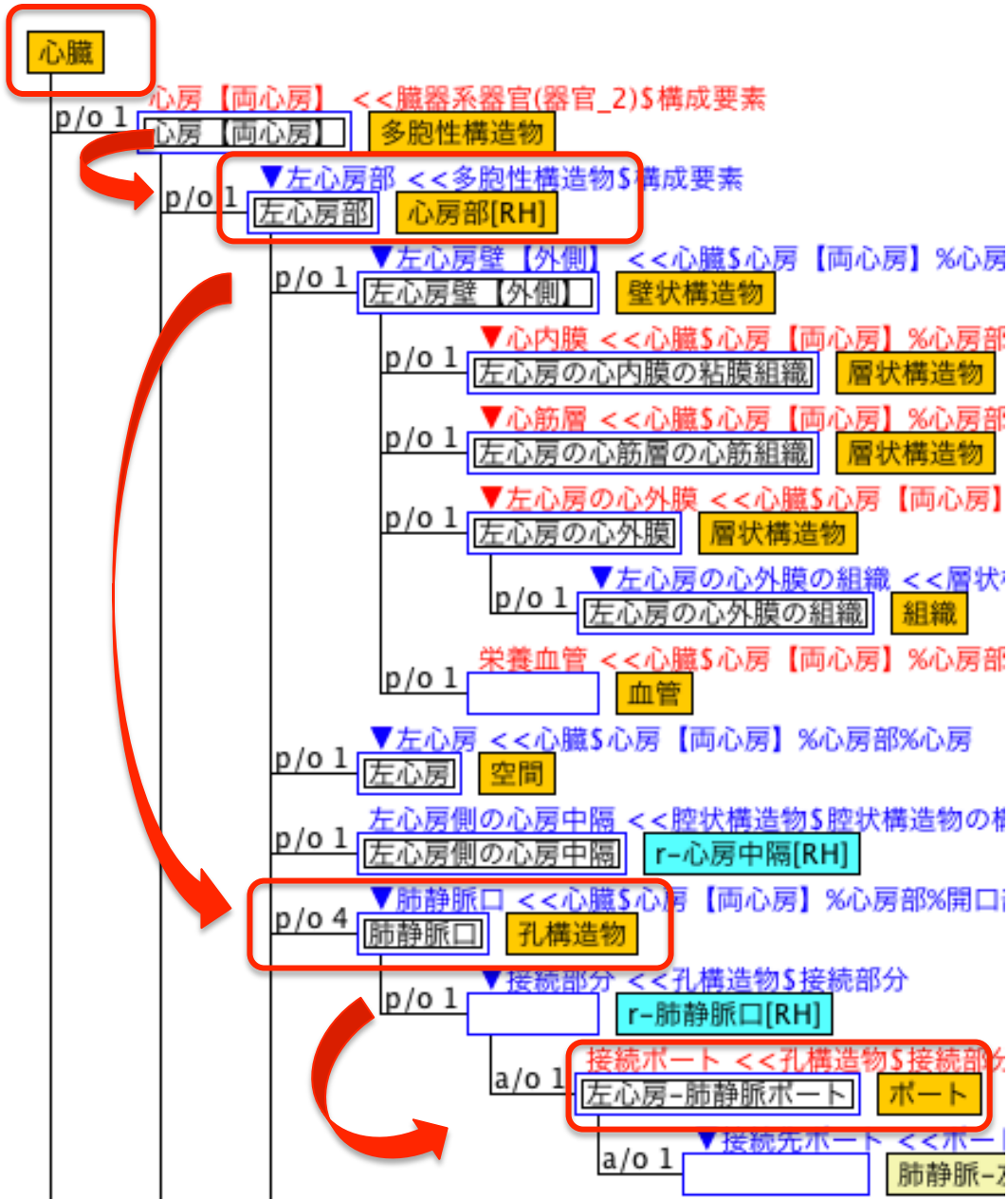


お互いのポートを相互参照

[接続ポートの独立化]
構造物の接続部分には「接続ポート」が仮想的に存在するものとし、ポートを介して相手と接続する



- ・接続部分のみを人工物に置換する手術操作を自然に記述
- ・「接続箇所の不具合」を「構造物の不具合」と分離して記述
(ex. 骨の脱臼・・・ 構造物同士は存在しているが“接続”が切れている)



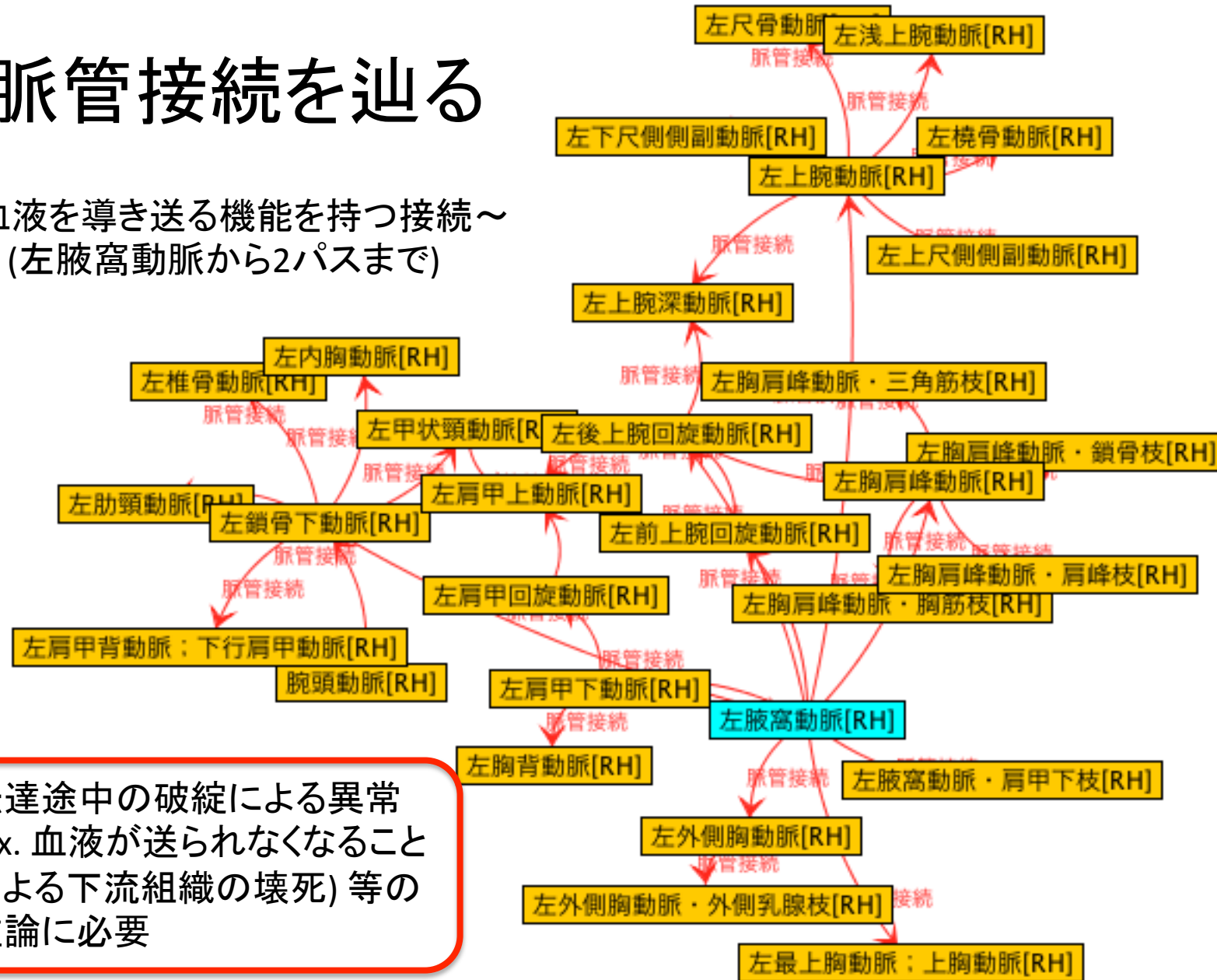
心臓での例

- ・「左心房部」の構成部分として【孔構造物】である「肺静脈口」が存在。
- ・それが持つ肺静脈への接続ポートを介して「肺静脈」へ接続
- ・複数の接続種類があり、伝達しているものが異なる

- 脈管接続
- 空間接続
- 毛細血管型接続
- 分布状接続
- 力学接続
- 接触接続

脈管接続を辿る

～血液を導き送る機能を持つ接続～
(左腋窩動脈から2パスまで)



伝達途中の破綻による異常
(ex. 血液が送られなくなるこ
による下流組織の壊死) 等の
推論に必要

解剖構造オントロジーの現状規模

- 5つのパートに分類して記述
 - 1) 器官, サブシステムなど抽象度の高い解剖構造や汎用的構造物
 - 2) 動脈系 (概念数: 6,400 接続数: 5,600 程度)
 - 3) 静脈系 (概念数: 3,600 接続数: 3,000 程度)
 - 4) 神経系 (概念数: 6,800 接続数: 3,200 程度)
 - 5) 筋骨格系 (骨・関節は共に 概念数: 200~250, 接続数: 350
筋肉・靭帯等は記述中)
- 主要な構造物は現時点でほぼ網羅
今後さらに粒度の細かい概念を追加拡充予定

臨床医学オントロジーの活用展望と課題

• 活用展望

説明能力！

– 診断・診療支援

- 患者状態からの診断候補群の提示
- 将来起こりうるリスクの提示, 適切な治療法の提示など

– 医学研究支援

- 自然言語で記載された診療情報の解析 (自然言語処理技術の高度化)
- クエリー拡張
ex) 重篤な疾患, 代謝系の異常, 関連症状・類似症例の検索など

– 教育

- 初学者の疾患概念理解の手助け (共通性と特殊性の理解促進)
- 定義を自然言語文変換することで、一般向け電子医学解説書生成

– 複数の施設間で、医学知識・診療情報をやりとりする際の「共通参照概念基盤」の提供 & 統計調査を支える基盤

- 標準化・医師の情報入力の際に患者状態に応じて適切な専門用語を提示
- 統計調査・情報記述の揺れの防止、国際疾病分類体系への自動分類

臨床医学オントロジーの活用展望と課題

• 今後の課題

– 十分に記述されていない情報の拡充

- 治療法 (医薬品・手術手技等)
- 異常状態の「観測」に関する知識 (検査値を元にした判断基準, どのような検査所見として身体内部の異常状態が観測されるか)

– 利活用アプリケーションの開発

- 実用的なアプリケーションからのフィードバックの収集

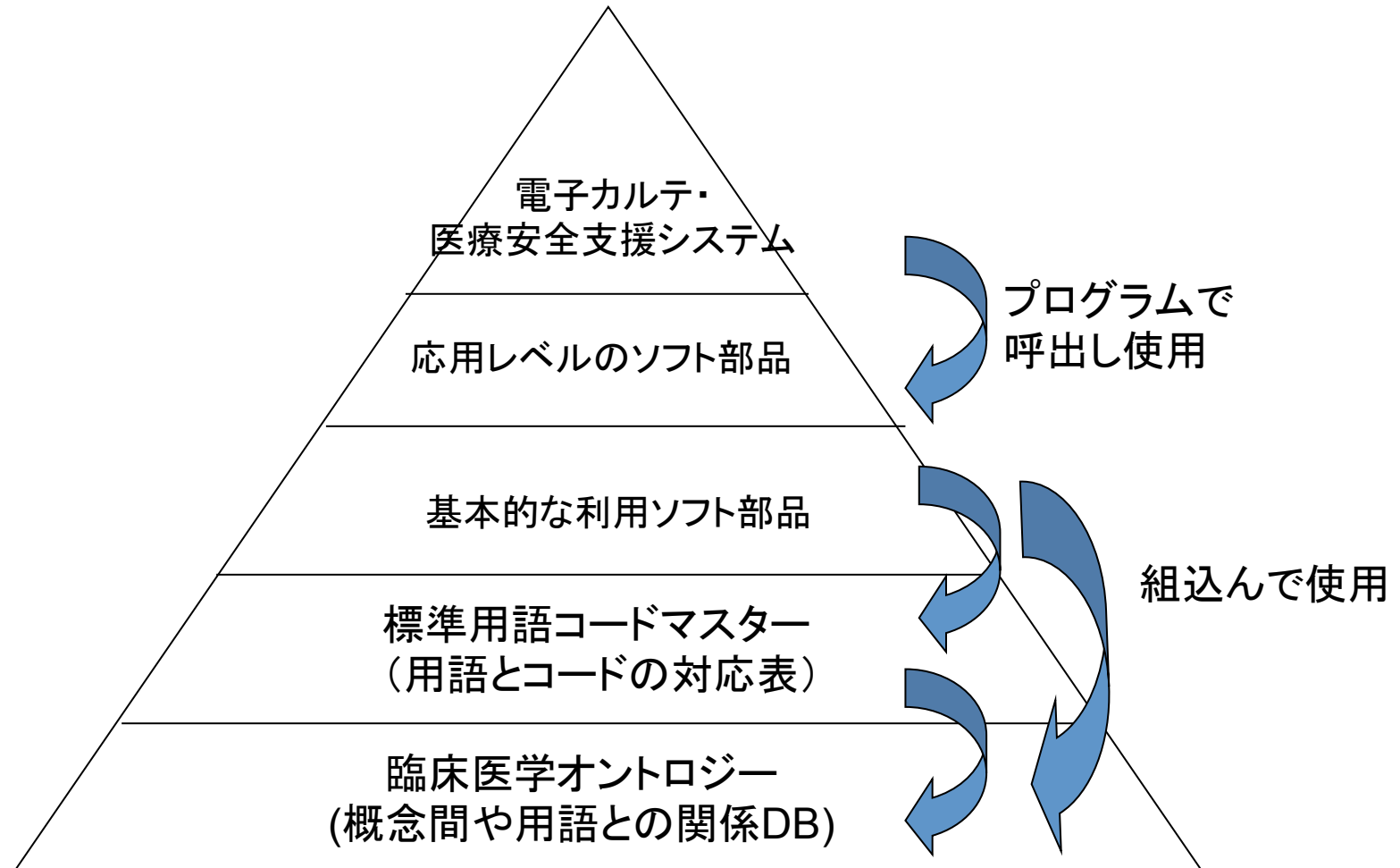
– 自然言語処理との親和性の拡充

- 診療情報中に現れる用語文字列 (ラベル) を大量に収集し、オントロジー中の概念との対応関係を付与

– メンテナンス体制と支援手法

- オントロジー内の知識をアップデートするための情報を、自然言語処理技術などの活用により半自動的に行う仕組み

臨床医学オントロジーとそのさまざまな応用レイヤー



まとめ

- 我が国における臨床医学オントロジー研究開発の概要、活用展望と課題
 - 各種の計算機による知識推論処理に親和性の高い記述フレームワーク
 - 知識処理にとどまらず、教育・共通参照概念基盤としての活用も見込まれる
 - 実用的なアプリケーションの開発も並行して行い、フィードバックを元にした継続的なコンテンツの改良が必要

(謝辞): 本研究は厚生労働省委託事業「医療情報システムのための医療知識基盤データベース研究開発事業」(2007-09, 2010-)により行われています