3.1.1 牛乳試験

- 牛乳が感染しているかもしれない。感染しているかどうか検知するのに、陽性か陰性かを示す試験がある。しかし試験は完璧ではない。感染していない牛乳に対し陽性を示すかもしれないし、感染している牛乳に対し陰性を示すかもしれない。

- 牛乳が感染しているか感染していないかの 2 つの仮定事象がある。2 つの事象が排他的で包括的であるため、これらは *yes* と *no* の状態を持つ変数 *Infected?*に属する。試験結果は陽性か陰性である。これに対し、*pos* と *neg* の状態を持つ変数 *Test* を作る。

- 2 つの変数の因果方向は *Infected?*から *Test* である（図 3.1）。

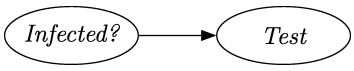

図 3.1 牛乳試験のベイジアンネットワーク

- 日に日に牛乳の状態は変わりうる。感染した牛乳を出す牛は時間とともに癒えるし、その日は大丈夫な牛でも次の日には感染している牛乳を出すリスクはある。今、農場主が試験を毎日行うとする。1 週間後には 7 日分の試験結果を得る。1 日ごとに図 3.1 のようなモデルがある。これら 7 つのモデルでは、過去の知識を現在の結論を得るためにつなげられるべきだ。前日の状態が当日の状態に影響させるのが自然である。これを表すのが図 3.2 である。

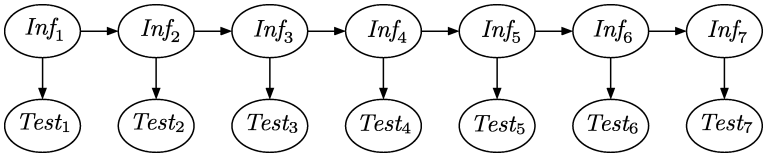

図 3.2 牛乳試験の 7 日間モデル

- 図 3.2 には d-分離から読めるいくつかの隠れた仮定がある。

- まず、モデルは *Markov property* を仮定している：今の状態が分かれば、過去は未来に対して何の影響も持たない。つまり、Inf_i が与えられた時、Inf_{i-1} は Inf_{i+1} から d-分離している。4 日目の牛乳が感染していることが分かっている時、5 日目の牛乳が感染しているかどうかを予測するのに使える。3 日目に感染していなかったことが分かっても、予測精度は上がらない。多くの病気に対しては、このような仮定は無効である。一部の病気は存

命期間がある。たとえば、今日インフルエンザにかかっているが、昨日は健康だったのであれば、明後日もおそらくインフルエンザにかかっているだろう。一方、四日間インフルエンザにかかっていたのであれば、明後日は治っている可能性は十分ある。もし図3．2のマルコフ特性が現実を反映していないのであれば、モデルは変えられるべきである。たとえば、さらに1日さかのぼる必要があるということになれば、図3.3のようになる。

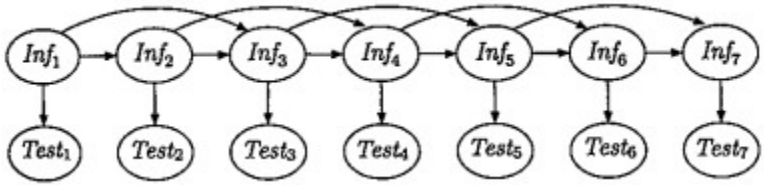

図3.3 二日間の感染期間がある7日間モデル

- 2つ目の隠された仮定は試験と関係がある。間にある感染ノードが与えられた時、どんな2つの試験ノードもd-分離されている。以前正しかったかどうかは試験が間違う確率とは無関係である。言い換えると、昨日試験が間違っていたことは今日試験が正しいかどうかに無関係である。もしこれが試験のふるまいを反映しないのであれば、図3.4のように昨日をモデルに組み込んでもよい。

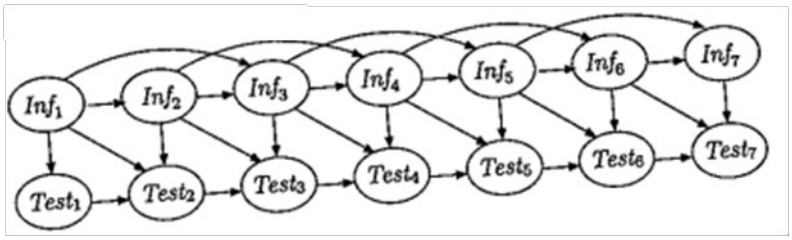

図3.4 二日間の感染期間と一日試験の精度期間のある7日間モデル

- **試験モデリングからの脱線**：どんな試験も完璧ではないと心得ることはよいことである。はっきりそうでないとわからない限り、試験の偽陽性と偽陰性の確率は与えられるべきだ。これがすべてではないが。間違った試験結果のメカニズムも考慮すべきである HIV 試験で偽陽性の確率が 10^{-5} とする。そして、ある人が陽性となったとする。今、再試験の選択肢があるが、これは役立つだろうか？試験が間違った結果を出すメカニズムによるだろう。HIV 感染しているかどうかにかかわらず、その人の血液が陽性を出す血液だったのであれば、再試験は無意味だ。一方、時々間違える試験だったのであれば、判決を下す前に再試験をするのは価値があるかもしれない（2度目が陰性だったのであれば、3回目を勧めてもよい）。図3.5にモデルを示す。

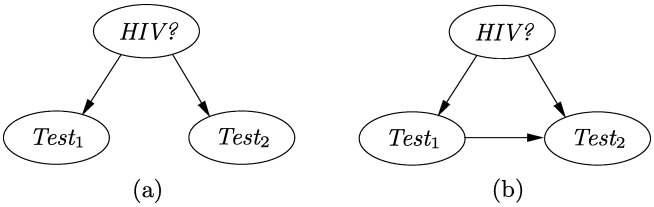

図 3.5(a)再試験が異なる結果を出しうるシナリオ
(b)再試験をしても同じ結果を生むシナリオ

3.1.2 風邪？それともアンギナ？

- 朝起きるとのどが痛い。風邪かもしれないし、アンギナ（のどの炎症）かもしれない。もっと洞察を得るために、温度を測ることやのどに黄点が付いていないか見ることができる。

- 5 つの仮定事象 *Cold?{no, yes}*と *Angina?{no, mild, severe}*がある。仮定事象はたがいに排他的で、包括的でなければならない。上記のような変数を用いてもよいかもしれないし *{no, cold, mild angina, severe angina}*を持つ変数 *Sick?*を用いてもよいかもしれない。後者の場合、両方を患う可能性が排除されている。2 つの変数 *Cold?*と *Angina?*を用いる。

- 情報変数は *Sore Throat?{no, yes}*、*See Spots{no, yes}*と *Fever?{no, low, high}*である。*Fever?*は連続型変数なので問題である。

- *Cold?* は *Sore Throat?*と *Fever?*に影響を与える一方、*Angina?*はすべてに影響を与える。モデルを図 3.6 に示す。

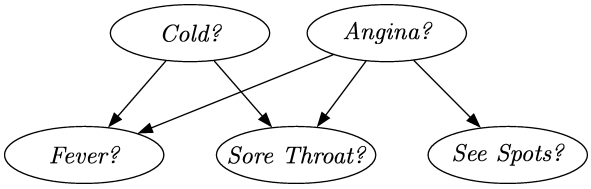

図 3.6 風邪？それともアンギナ？のモデル